

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE DERECHO



TESIS DOCTORAL

Bancos de datos jurídicos.

Pasado, presente y futuro

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Teresa Genoveva Vargas Osorno

Director

Emilio Suñé Llinás

Madrid, 2015

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE DERECHO



BANCOS DE DATOS JURÍDICOS
PASADO, PRESENTE Y FUTURO

TESIS DOCTORAL PRESENTADA POR

Teresa Genoveva Vargas Osorno

DIRECTOR: DR. EMILIO SUÑÉ LLINÁS

MADRID 2015

BANCOS DE DATOS JURÍDICOS
PASADO, PRESENTE Y FUTURO

CONTENIDO

LISTA DE FIGURAS	5
JUSTIFICACION	6
1. BANCOS DE DATOS JURÍDICOS: PASADO.....	13
1.1. Generalidades	13
1.2. Concepto.....	20
1.3. Antecedentes	23
1.4. Clasificación	27
1.5. Sistemas De Tratamiento De La Información Jurídica	28
1.5.1. Sistema referencial.....	29
1.5.2. Sistema full-text.....	30
1.5.3. Abstract jurídico	32
1.5.4. Método de indexación	33
1.6. Instrumentos lingüísticos.....	35
1.6.1. Tesoros.....	36
1.6.1.1. Clasificación de los tesauros.....	39
1.6.1.1.1. Según su utilización:.....	39
1.6.1.1.2. Según su idioma	44
1.6.1.1.3. Según su estructura.....	45
1.6.1.1.4. Según su contenido.....	46
1.6.2. Léxico	51
1.7 . Recuperación de la Información.....	57
1.7.1. Lógica formal.....	58
1.7.2. Operadores lógicos	62
1.7.3. Operadores sintácticos	65
1.7.4. Operadores de extensión.....	67
1.8. Elaboración de bancos de datos jurídicos.....	68
1.9. Panorama internacional	84
1.9.1. Unión Europea.....	84
1.9.2. Bélgica.....	86
1.9.3. Canadá.....	87
1.9.4. España.....	88
1.9.5. Estados Unidos	89
1.9.6. Italia.....	92
1.9.7. Francia.....	94
1.9.8. Alemania Federal.....	96
1.9.9. Brasil.....	96
1.9.10. Reino Unido	97

1.9.11.	<i>Perú</i>	98
1.9.12.	<i>Venezuela</i>	98
1.9.13.	<i>México</i>	99
1.9.14.	<i>Colombia</i>	100
2.	BANCOS DE DATOS JURÍDICOS: PRESENTE	110
2.1.	Introducción.....	110
2.2.	Historia de Internet	113
2.3.	La Incidencia de la World Wide Web en las Bases De Datos.....	119
2.4.	Web 1.0	126
2.5.	Web 2.0	134
2.6.	Antecedentes de los Search Engine generales	143
2.7.	Concepto de Buscador	155
2.8.	Funcionamiento de los buscadores	161
2.9.	Límites de los buscadores.....	166
2.10.	Search Engines mejor posicionados en la actualidad. Modelo De Google.	169
2.11.	Problemas que suscitan los buscadores o search engines en la actualidad	194
2.12.	Buscadores temáticos	211
2.13.	Buscadores especializados	214
2.14.	Antecedentes de los sistemas de recuperación de información jurídica.....	217
2.15.	Bancos de datos jurídicos en Internet	224
2.15.1.	<i>Bélgica</i>	226
2.15.2.	<i>Francia</i>	228
2.15.3.	<i>Reino Unido</i>	231
2.15.4.	<i>Canadá</i>	235
2.15.5.	<i>España</i>	238
2.15.6.	<i>Unión Europea</i>	246
2.15.7.	<i>Alemania</i>	252
2.15.8.	<i>Venezuela</i>	255
2.15.9.	<i>Colombia</i>	259
2.15.10.	<i>Dinamarca</i>	263
2.15.11.	<i>Países Bajos</i>	263
2.15.12.	<i>Estados Unidos</i>	264
2.15.13.	<i>Italia</i>	274
2.15.14.	<i>Brasil</i>	279
2.15.15.	<i>Argentina</i>	281
2.15.16.	<i>Ecuador</i>	292
2.15.17.	<i>México</i>	295

2.15.18. Perú.....	300
2.15.19. Costa Rica.....	301
2.15.20. Puerto Rico.....	303
2.15.21. Uruguay.....	306
2.15.22. Panamá.....	308
2.15.23. África Del Oeste.....	309
3. BANCO DE DATOS JURÍDICOS: FUTURO	311
3.1. La era de la sobreinformación y el Big Data.....	311
3.2. El análisis de los datos jurídicos en la era del Big Data.....	317
3.3. La era de la web 3.0 o Web Semántica	318
3.3.1. Web Semántica y su aplicación en derecho	327
3.3.2. Herramientas de la Web Semántica.....	335
3.3.3. Metadatos en la Web Semántica - importancia en la informática jurídica	337
3.4. Importancia del uso de vocabulario semántico en los metadatos: las ontologías	341
3.5. ¿Cómo se genera inteligencia a partir de los datos jurídicos?.....	343
3.5.1. Inteligencia artificial.....	344
3.5.2. La inteligencia jurídica artificial.....	348
3.6. Hacia la creación de una ontología jurídica	355
3.6.1. El ordenamiento Jurídico como un sistema complejo	355
3.6.2. Definición de la ontología jurídica.....	360
3.6.3. La creación de una ontología jurídica.....	368
3.7. Ordenamiento jurídico: sistema complejo con uso de lenguaje natural.....	369
3.7.1. Tratamiento del ordenamiento jurídico como sistema complejo.....	376
3.8. Hacia una concepción de las estructuras arbóreas.....	380
3.9. La aplicación de la probabilidad a los conceptos ontológicos	394
3.10. Formalización de la ontología	395
3.11. La materialización semántica de la ontología	397
CONCLUSIONES.....	400
BIBLIOGRAFÍA.....	407
BIBLIOGRAPHY	436
ABSTRACT	437

LISTA DE FIGURAS

Figura 1. Macrotesauro en derecho infomático.....	73
Figura 2. Representación de la web 1.0	134
Figura 3. Representación de la web 2.0	143
Figura 4. Representación gráfica de inteligencia jurídica artificial	354
Figura 5. Ordenamiento jurídico como sistema complejo	357
Figura 6. Principios Diferenciales en DOE. (Differential Ontology Editor)	375
Figura 7. Informaciones Lingüísticas en DOE. (Differential Ontology Editor)	375
Figura 8. Estructura arbórea de artículos Código Civil Colombiano.....	384
Figura 9. Estructura arbórea de apartes del Código Civil Colombiano	387
Figura 10. Análisis conjunto de convergencia de estructuras arboreas	388

JUSTIFICACION

Para estudiar la actualidad jurídica, es menester extrapolar la definición, los antecedentes y las diferentes percepciones históricas que se pueden observar en torno a los bancos de datos jurídicos. Serán analizados desde la perspectiva de la Informática Jurídica, que se presenta como aquella especialidad del Derecho, encargada del estudio de la aplicación de instrumentos electrónicos e informáticos, en la resolución de problemas jurídicos. En esta misma línea, el presente, pasado y futuro de los bancos, debe ser analizado a profundidad, por cuanto, para establecer y enervar la importancia de los mismos, no se puede eludir un estudio pormenorizado de sus orígenes y de las proyecciones que, tanto el Derecho como otras disciplinas que intervienen en su desarrollo, han sabido establecer.

La imprenta hizo que el conocimiento fuera, no sólo fácil de preservar sino accesible a quien lo deseara. Un libro, ya no era un objeto de lujo en manos de unos pocos. El costo de fabricarlo había bajado, también porque había un aumento en la producción de papel, el cual se hacía con ropa vieja. Luego, el mundo europeo se vio sacudido por un evento de locura: el descubrimiento de América. Las cartas de Colón y Vespucio, divulgaron el conocimiento como no hubiera sido posible antes de la imprenta de Gutenberg. Del mismo modo se generalizaron escritos que revolucionaban la manera tradicional de pensar. Todo iba quedando al alcance de todos: lo seguro, lo peligroso, lo nuevo, lo antiguo.

A veces parece que quisiéramos imitar a Funes, el memorioso, aquel personaje de Jorge Luis Borges, que se cayó de un caballo a los 19 años y, como

consecuencia del golpe, adquirió una memoria tan prodigiosa que él mismo decía: “Más recuerdos tengo yo que los que habrán tenido todos los hombres desde que el mundo es mundo.” Y también decía: “Mi memoria es como vaciadero de basuras.”

Funes carecía de la capacidad del pensamiento: no olvidaba pero no abstraía, no generalizaba. Una memoria así, libre de discernimiento como nudos inertes en un quipu mecánico, abierta a quien interprete los datos según sus propias necesidades e idiosincrasia, ¿puede llamarse un basurero? ¿No es más bien una mina de riqueza inagotable?

Tal vez, los humanos no quisieran parecerse a Funes, pero sí crearon algo que hace y logran lo mismo que aquellos: los Bancos de Datos.

La palabra “información” está compuesta de dos partes: “in” como prefijo que significa dirigir hacia adentro, y, “formatio” que, en latín, se traduce como dar forma a algo que viene de afuera. ¿Y qué es ese algo? Un montón de datos, ni más ni menos, que agrupados y organizados conforman lo que se ha dado en llamar Ciencia de la Información y que, en el lenguaje moderno de la Informática, se denomina *Base de Datos*. Es como un preámbulo para enviar un mensaje, capaz de incidir en la conciencia de quién lo recibe y altera o modifica su comportamiento y le ayuda a decidir. Y si se refiere al tema jurídico, los bancos de datos (con el nombre del dato que agrupe y organice), tienen la función principal de estar disponibles en el computador, valga decir, en Internet, y de poderse consultar el tiempo y las veces que el usuario quiera. Es una información que se convierte en conocimiento, por

medio del cual, el ser humano da sentido a las cosas, a los seres y a su entorno. Aquí, se puede afirmar que, sabio es aquel que mejor emplea lo que sabe, lo que conoce.

Desde la década de 1980, Internet comenzó a expandirse por todo el mundo. William Gibson creó el término “cibespacio”, un espacio sin fronteras territoriales, con un lenguaje común. Abrió nuevos horizontes para la educación. Los universitarios podían participar, en las clases dictadas por insignes catedráticos, en otros países. Y cualquier persona, desde un estudiante hasta un ama de casa, podía informarse sobre cualquier tema. “San Google” se convirtió en una figura cotidiana, de devoción familiar.

Se ha dicho que internet es una biblioteca mundial. Ya no volveremos a llorar, metafóricamente hablando, por el incendio de la biblioteca de Alejandría, donde se perdió gran parte del conocimiento humano. Ya la biblioteca está en todas las ciudades, en todos los hogares, en la habitación propia, por lo que cada ser humano puede reflexionar y tomar las propias decisiones de su vida y destino.

Y el gran resultado, después de esta gran innovación, va a ser la transformación de la humanidad en sus relaciones y en el respeto por el otro, en medir las consecuencias de las propias acciones y en modificar su propia actitud frente al mundo, a los seres y las cosas.

Con el advenimiento de la denominada *sociedad de la información*¹ y de las consecuencias, casi obvias, que de suyo se desprenden, el mundo entero afronta nuevos retos y, especialmente, los profesionales del Derecho, quienes se verán abocados a combatir la acreencia social que se deriva del uso de las tecnologías, mediante la producción e implementación del conocimiento jurídico.

En este nuevo contexto social, la información, por encima de todo, se pondera como el producto, por excelencia, en casi todas las esferas en las que el hombre tiene algún tipo de injerencia. El Derecho, por supuesto, no podía ser la excepción.

La producción de material jurídico (decretos, leyes, jurisprudencia, doctrina, etc.), representa un arsenal invaluable para que, tanto el abogado como cualquier ciudadano, en general, refuerce sus derechos bajo un manto especializado de protección, en el que el acceso, dominio e interpretación de la información, se convierta en la llave a utilizar para poder abrir las cerraduras de la tutela del Estado.

No obstante, a la par con el surgimiento de las llamadas tecnologías de la información y las comunicaciones, surgen también obstáculos que dificultan la democratización de la información. La bien llamada *brecha digital*, constituye un categórico para todos aquellos escenarios en el que, el acceso, por parte de la sociedad y en términos de igualdad, a las tecnologías, se ve truncado por diferentes razones, entre ellas, la calidad y la eficiencia en la gestión de la información.

¹ “La sociedad de la información es un estadio de desarrollo social caracterizado por la capacidad de sus miembros (ciudadanos, empresas y administración pública) para obtener y compartir cualquier información, instantáneamente, desde cualquier lugar y en la forma que se prefiera”. DE LA FUENTE, F. (2004). *Los sistemas de la información en la sociedad del conocimiento*. Madrid. ESIC Editorial. p. 2.

En este sentido, cabe perfectamente el objeto principal de este escrito que no es otro que el de presentar, al lector, una visión descriptiva sobre los bancos de datos jurídicos y exaltarlos como una respuesta a la problemática de la brecha digital jurídica, en todo el mundo.

Este trabajo también tiene como propósito, presentar un marco global sobre la aplicación geográfica de los bancos de datos jurídicos, a fin de comparar, entre otras cosas, los fundamentos prácticos que tienen influencia, tanto en particulares como en entes gubernamentales, a la hora de administrar, conservar y divulgar la información jurídico-legal, en atención a las facilidades y restricciones que la población, en general, tiene en su acceso. En Colombia, por ejemplo, desde la década de los 70's, se ha impulsado la creación de diferentes modelos de bancos de datos jurídicos. De la mano del Gobierno Nacional y con el impulso de diferentes agentes privados, como empresas y academias, la ardua labor de recopilación y democratización de la información jurídica, ha obtenido importantes resultados y, hoy por hoy, continúa su curso hacia la democratización absoluta del *dato* jurídico.

De forma más concreta, este estudio busca:

- **Destacar la implementación de las nuevas tecnologías de la información y la comunicación (TIC) en el ejercicio de la abogacía.** Es de saber común que, actualmente, el desempeño de la profesión del abogado no se concreta estrictamente en la lectura aislada de códigos, leyes, jurisprudencia y doctrina, documentada en formato físico; más bien, se hace predominante la inclusión del grueso de los documentos jurídicos, bajo la noción de un “conjunto”,

denominado *ordenamiento jurídico* que, a su vez, con la aplicación de las TIC, puede ser igualmente analizado como un todo.

- **Presentar los bancos de datos jurídicos como una herramienta idónea para la protección de los derechos de las personas, en general.** Toda vez que, si bien la protección que se deriva de la publicidad y divulgación de la información jurídica (normas, jurisprudencia, etc.) trae consigo un ámbito de “amparo” determinado, aún hay instrumentos, como los bancos de datos, que procuran la tutela de los derechos de las personas, en un nivel mucho más amplio, general y pluralista. Asimismo, dan la oportunidad, al usuario, de encontrar una información organizada, pues, esta es presentada, en forma ordenada, permitiendo una consulta rápida, ágil y veraz.
- **Estudiar en el contexto, características, fundamentos y aplicación de la *Web Semántica*, como paso ineludible en la gestión de información jurídica.** La categorización masiva de información y su vinculación, mediante las diversas plataformas de búsqueda, en internet, actualmente se manifiesta bajo la tendencia de incluir la semántica en sus búsquedas para, de esta forma, construir niveles de sucesiva abstracción y control, que permitan realizar un filtrado automático de datos estructurados, para que los usuarios de los buscadores encuentren, como resultado, no una información cualquiera sino una que tenga un contexto y diga, incluso, más que su contenido mismo.

Podemos afirmar que, hasta ahora, ha sido un monopolio del Estado, de las universidades y de pocos, el conocimiento de las normas jurídicas y su aplicación en

sentencias y doctrinas, lo que de alguna manera contraría la obligación general de conocer la ley, los procedimientos, las excepciones, etc. Siempre se ha dicho que nadie puede excusarse en la ignorancia de la ley, para eludir su aplicación, ya que la ley se reputa conocida por todos. Pero es, en este momento, cuando verdaderamente puede ser realidad tal premisa jurídica. Ese despliegue en la red, de los bancos de datos jurídicos, con el trabajo de la semántica, se convierte en una herramienta eficaz, no sólo para el juzgador sino también para el procesado, para el que defiende, para el que acusa y para unir el pensamiento universal jurídico, que cada día debe propender más, hacía una Justicia Universal.

Estamos en el momento justo cuando la web nos facilita la conjugación real del derecho de igualdad, consagrado como derecho fundamental en nuestra Constitución porque aquella permite que, sin distinciones de ninguna clase, todo el mundo pueda acceder a un sin número de conocimientos. Y nuestra tesis busca, ante aquello, presentar una concepción nueva, en los tópicos tratados, como motivar el análisis de los medios para que aquellos sean más perfectos y puedan acceder todas las personas que los necesiten, dándoles una útil entrega como fecunda enseñanza.

1. BANCOS DE DATOS JURÍDICOS: PASADO

1.1. Generalidades

La época que nos correspondió vivir, está marcada por la revolución científica. Nos encontramos ante la información automatizada, puesta en el primer plano del horizonte de la llamada sociedad tecnológica, con una serie de consecuencias, de distinto orden (social, económico, jurídico, cultural, político etc.), que ella comporta y que ha puesto las bases para la creación de una civilización planetaria homogénea, pues, por primera vez en la historia del hombre, el conocimiento ha pasado a ser patrimonio de la humanidad, en su conjunto; sin barreras ni limitaciones de especie alguna, contribuyendo, de otra parte, a crear sociedades globalizadas y con acceso a la información de manera indiscriminada, efectiva y real. Quien posee la información pueden decidir con pluralidad de opciones. ¿Nos estamos acercando, acaso, al ideal buscado durante siglos por la humanidad, relacionado con los principios de igualdad y educación?

La denominada sociedad de la información, también llamada sociedad del conocimiento, o sociedad postmoderna, ha generado que la creación, distribución y manipulación de la información cobre importancia dentro de aspectos culturales y económicos. Esta dimensión del conocimiento, no le es ajena al mundo jurídico, razón por la cual surge el concepto de derecho de la informática e informática jurídica.

De manera que, encontramos que la producción y distribución del conocimiento, la cantidad de empleos que se basan en la manipulación y manejo de información, actualmente, es mayor que los relacionados con algún tipo de esfuerzo físico, de ahí que se inicia la transformación de la sociedad industrial a una sociedad de la información, que se traduce en un incremento de empleos dentro del sector servicios y una disminución dentro del sector industrial, y como consecuencia se le atribuye una mayor importancia a los bienes intangibles.

Dentro de la sociedad de la información, existen unos requisitos *sine qua non* para su existencia; por ello encontramos, en primera instancia, como objeto y elemento de análisis *la información*, que se encuentra hoy por hoy circulando de manera indiscriminada por la red. Por otra parte, se encuentran los *usuarios* de la información anteriormente mencionada, que para efectos de acceder a la misma utilizan la *infraestructura* pre establecida para estos fines.

En dicho modelo de sociedad, toma importancia el concepto de las denominadas Tecnologías de la Información y la Comunicación (TIC), que construyen nuevas formas de desarrollo informático. No obstante, como consecuencia de esto, se genera una dependencia tecnológica de parte de la sociedad, quienes se ven afectados dentro de los diferentes procesos de desarrollo al no contar con los elementos informáticos necesarios para ejecutar sus funciones.

Ahora bien, pese a las ventajas que ofrece la sociedad de la información, uno de los problemas que se presenta como un obstáculo, es el fenómeno de *la brecha*

*digital*², que se puede presentar entre los grupos sociales que cuentan con acceso a medios informáticos y de telecomunicaciones y los que carecen de ella. Así mismo, dentro de los grupos que acceden a los medios informáticos y de telecomunicaciones, se pueden diferenciar quiénes poseen el conocimiento adecuado para organizar, seleccionar y discriminar la información y aquellos que no lo poseen.

Conforme a lo anterior, por medios estadísticos, la CIA (Central Intelligence Agency), dentro de un estudio realizado, determinó que Colombia se encuentra en el puesto número 21 a nivel mundial con 17.117 usuarios que contaban para el año 2008 con acceso a internet, estas estadísticas varían de país a país y pueden incluir usuarios que acceden a internet diferentes veces a la semana o que acceden solo una vez en un periodo de varios meses³.

En lo relacionado con las centrales de internet, el mismo estudio, determinó que Colombia a 2009, contaba con 2.217 millones, ubicándose en el puesto 30, esta es una lista de número de administradores de internet, un *internet host* es un computador conectado directamente a internet, normalmente un proveedor de servicios de internet o *internet service provider* (ISP)⁴.

² La brecha digital, es un concepto que surge a raíz de la denominada sociedad de la información, se constituye como un obstáculo al acceso de la misma, a causa de la desigualdad socioeconómica entre las comunidades que tienen acceso a internet y las que carecen de este servicio, al igual, hacen referencia a la imposibilidad de acceso al computador personal, la telefonía móvil y la banda ancha.

³ AGENCIA CENTRAL DE INTELIGENCIA. CIA. (2013). The World Factbook. Washington, DC: Central Intelligence Agency. <https://www.cia.gov/library/publications/the-world-factbook/rankorder/2153rank.html?countryName=Colombia&countryCode=co®ionCode=sa&rank=21#co>

⁴ Ibídem

Según Internet World Statu, actualizado en el 2005, de los 985 millones de internautas conectados, casi el 70% vive en los países industrializados, donde reside el 15% de la población mundial. Mientras que Europa y Estados Unidos suman quinientos millones de usuarios, en todo el continente africano no hay más que cuatro millones, y estas diferencias se manifiestan así mismo entre hombres y mujeres, ciudad o campo, edades, estatus sociales, paralelamente a las « brechas » de siempre: el acceso a la sanidad, a la educación, la mortalidad infantil, el hambre, la pobreza. Datos de 2006 de la Unión Internacional de Telecomunicaciones muestran que mientras el 58.6% de los habitantes de los países desarrollados tienen acceso a Internet, en los países en vías de desarrollo apenas el 10.2% de los habitantes tiene acceso a esta tecnología⁵.

Observamos entonces, que dentro de la sociedad de la información, la brecha digital se constituye como un problema del cual no podemos ser ajenos, en especial, si se tiene presente que las cifras que se presentan son desfavorables para Colombia, por esta razón se hace necesario que dentro de la ciencia jurídica que nos ocupa, se propugne por el establecimiento eficaz de los derechos de comunicación dentro de la sociedad de la información, observando a las tecnologías de la información y la comunicación, no como un fin sino como un medio para impulsar un cambio social.

Es de anotar, que una de las causas de la brecha digital, es la calidad, organización y discriminación de la información, aspecto que se presenta como un hecho posterior al acceso de la información, que genera brechas dentro de los grupos sociales debido al desconocimiento del manejo de la información. Este es uno de los objetos del

⁵ *Brecha digital*. (2015). Consultada en enero 25, 2015.
http://es.wikipedia.org/w/index.php?title=Brecha_digital&oldid=79306248.

presente estudio, establecer las causas y las soluciones sobre las diferencias que se presentan una vez se tiene acceso a la información, el acceso a usos avanzados o técnicos, los centros de inteligencia artificial que permitan relacionar la información almacenada, analizarla y sacar conclusiones lógicas, la ontología informática que clasifica y realiza un razonamiento inductivo⁶ de la información.

En los albores de una nueva era denominada *sociedad del conocimiento*, se ha originado una revolución industrial: la de las nuevas tecnologías que conlleva un nuevo avance de la humanidad donde el compartir los conocimientos a través de la información, es el objetivo principal, tal como se señala, en el informe mundial de la UNESCO, «Hacia las Sociedades del Conocimiento» publicado con motivo de la Cumbre Mundial de Túnez donde uno de los temas fue De la Sociedad de la Información a la sociedad del conocimiento”⁷.

Dentro de ese informe principalmente, menciona la UNESCO que:

“La sociedad mundial de la información en gestación sólo cobrará su verdadero sentido si se convierte en un medio al servicio de un fin más elevado y deseable: La construcción a nivel mundial de sociedades del conocimiento que sean fuentes de desarrollo para todos, y sobre todo para los países menos adelantados. Para lograrlo, dos desafíos planteados por la

⁶ Razonamiento inductivo, es aquel pensamiento lógico que se obtiene a partir de premisas con un contenido particular. Dentro de la ontología informática, el razonamiento inductivo es más óptimo a mayor cantidad de datos, por cuanto aumenta la probabilidad de un correcto razonamiento.

⁷ *Hacia las sociedades del conocimiento: Informe Mundial de la UNESCO*. Dirigido por BINDE, Jérôme. UNESCO. Paris. <http://unesdoc.unesco.org/images/0013/001321/132114f.pdf>. 2005. P. 29.

revolución de la información revisten una importancia particular: El acceso a la información para todos y el futuro de la libertad de expresión⁸”.

Como bien lo menciona el informe de la UNESCO, las sociedades del conocimiento deben constituirse como fuentes de desarrollo, “Un elemento central de las sociedades del conocimiento es la “capacidad para identificar, producir, tratar, transformar, difundir y utilizar la información con vistas a crear y aplicar los conocimientos necesarios para el desarrollo humano”⁹.

En este punto es relevante mencionar que:

“Según Castells (2000), hemos pasado de una sociedad industrial a una posindustrial cuya nueva base se sustenta en el manejo y la transmisión de información de toda clase, esto es, la sociedad de la información, en la que la acumulación de productos se ha sustituido por la acumulación de conocimiento. Así, se entiende que el nuevo reto de la Unión Europea fijado en la Cumbre de Lisboa de 2000 sea convertirse en 2010 en la sociedad del conocimiento más competitiva del mundo. Pues bien, en esta nueva sociedad, la información legal emerge como un pilar fundamental, ocupando un lugar prioritario, dentro de los desarrollos de e-gobierno”¹⁰.

⁸ Ibídem, p. 29. “Además, ¿Cómo se puede hablar de sociedad mundial de la información, cuando la libre circulación de información se ve obstaculizada o cuando la propia información es objeto de censuras y manipulaciones?”

⁹ Ibídem.

¹⁰ ALVITE DÍEZ, M. L. (2009, Enero-Junio). Las bases de datos jurídicas y el uso del lenguaje XML en España. Scire, 15(1), 2. P.35 <http://www.iberid.eu/ojs/index.php/scire/article/view/1768/1739>

Ahora bien, el conocimiento debe contener dos características importantes: La no exclusividad, porque el conocimiento debe ser fuente de desarrollo y para que esto se logre, es necesario que se constituya como un bien común, la otra característica del conocimiento es la no rivalidad, de manera que no se puede impedir el uso del conocimiento. Compartir los conocimientos será la vía que nos permitirá luchar eficazmente contra la pobreza y promover un desarrollo humano sostenible. Nuestra lucha deberá estar dirigida a vencer los obstáculos que se oponen al advenimiento de las sociedades del conocimiento compartido.

De otra parte, de las relaciones entre Informática y Derecho surge el derecho informático, en donde encontramos dos extremos: En un extremo, la informática al servicio del derecho, que implica una incursión en la informática jurídica y, en el otro extremo, si se considera a la informática en todos sus aspectos como objeto del derecho, estaremos frente al campo del Derecho de la Informática. Podemos decir, en síntesis, que así como en la Informática Jurídica, es la informática la que está al servicio del derecho, en el caso del derecho de la informática, los papeles se invierten.

Una vez mencionado lo anterior, el tema que nos ocupa es la tecnología informática aplicada a la sistematización y consolidación del Derecho, también denominada *informática jurídica*. Partiremos entonces, con el estudio de los aspectos generales, particularmente su concepto, antecedentes, clasificación, concentrándonos en el desarrollo de la informática jurídica documental, que es el punto central de esta tesis doctoral.

1.2. Concepto

La disciplina denominada informática jurídica, que forma parte de la antes mencionada sociedad de la información, principalmente busca otorgar herramientas útiles a la profesión jurídica, mediante las cuales se pueda generar la sistematización y organización del ordenamiento jurídico con el único fin de crear un sistema de información jurídica que otorgue una visión globalizada del conocimiento de manera que no quede excluida ningún aporte que pueda llegar a ser relevante dentro de la ejecución y búsqueda de los diferentes conceptos o temáticas con las cuales cuenta el derecho.

Múltiples han sido las definiciones que se han dado sobre el concepto de *informática jurídica*, según Antonio Enrique Pérez Luño la informática jurídica se define como el

“tratamiento automatizado de: las fuentes de conocimiento jurídico a través de los sistemas de documentación legislativa, jurisprudencial y doctrinal «informática jurídica documental», de las fuentes de producción jurídica, a través de la elaboración informática de los factores lógico-formales que concurren en proceso legislativo y en la decisión judicial «informática jurídica decisional »; y los procesos de organización de la infraestructura o medios instrumentales con los que se gestiona el derecho «informática jurídica de gestión »¹¹”

¹¹ PÉREZ LUÑO, A. E. (1996). Manual de informática y derecho. Barcelona: Ariel. p. 22

Por su parte, José Manuel García García, define la informática jurídica, como “aquella ciencia que tiene por objeto el tratamiento automático de la documentación jurídica con fines de información”¹².

En otras palabras, por informática jurídica debe entenderse el conjunto de estudios e instrumentos derivados de la aplicación de la informática al derecho, o más precisamente, a los procesos de creación, aplicación y conocimiento del derecho¹³.

Juan Diego Castro Fernández, en su libro “*Juristas y Computadoras*”, la define como aquella disciplina “a la que corresponde el empleo de los métodos y las técnicas del procesamiento de datos en el arte y en la ciencia del derecho”¹⁴.

El doctrinante, Emilio Suñé Llinás, ha entendido por informática jurídica, “la aplicación de los ordenadores electrónicos dirigida a la resolución de problemas jurídicos.”, y afirma que “la informática jurídica está orientada a la resolución de problemas jurídicos, porque aunque ello fuese factible, entiende que jamás se debe encomendar a un ordenador la decisión de un problema jurídico.”¹⁵. Suñé deje claro que la informática jurídica ha de limitarse a servir de apoyo a los distintos profesionales del derecho para que éstos puedan llevar a cabo de forma eficaz la labor que les es propia e inexcusable. Por tanto, la informática que se ponga a

¹² GARCÍA GARCÍA, J. (1979). *Problemática de la documentación jurídica*. Ponencia en el III curso de informática jurídica para juristas hispanoamericanos.

¹³ FIX FIERRO, H. (1990). *Informática y documentación jurídica*. México: Universidad Nacional Autónoma de México. UNAM. p. 46.

¹⁴ CASTRO FERNÁNDEZ, J. (1992). *Juristas y computadoras*. San José de Costarica: Juriscomputación S.A. P. 23.

¹⁵ SUÑÉ LLINÁS, E. (1986). *Introducción a la informática jurídica y al derecho de la informática*. Revista de la Facultad de Derecho de la Universidad Complutense de Madrid. P. 65.

disposición de los juristas solo será jurídica cuando resuelva cuestiones propias de la profesión¹⁶.

Entre el Derecho y la Informática, se podrían apreciar dos tipos de interrelaciones. Si se toma como enfoque el aspecto netamente instrumental, se está haciendo referencia a la *informática jurídica*, es decir, la informática al servicio del derecho. Pero al considerar a la informática como objeto del Derecho, se hace alusión al *Derecho Informático*.

En ese orden de ideas, siendo la *informática jurídica* el tema que nos ocupa, podemos definirla como aquella ciencia que estudia la utilización de los computadores o elementos electrónicos como herramienta en el derecho; es decir, la ayuda que este uso presta al desarrollo y aplicación dentro del ámbito jurídico. En otras palabras, podemos definir la *informática jurídica*, como la aplicación de las herramientas informáticas a la ciencia del derecho, cuyo objeto es el tratamiento automático de la información jurídica, para facilitar al abogado, la búsqueda fácil y ordenada de la misma.

En este sentido, es importante hacer claridad, que actualmente no solamente se considera como la aplicación de herramientas informáticas al derecho, sino como la aplicación del conjunto de medios electrónicos. La Informática Jurídica es básicamente una técnica en el tratamiento de información enfocada al derecho, por lo tanto al igual que ésta es de carácter jurídico bien podría ser de otra índole, del

¹⁶ Ibídem.

mismo modo que se aplica a la medicina, la economía, la contabilidad o cualquiera otra materia.

Para concluir, tenemos entonces que la *informática jurídica*, es la disciplina que aplica la tecnología informática y electrónica a la ciencia del derecho, constituida por los procedimientos para sistematizar bases de datos jurídicas.

1.3. Antecedentes

El origen de la *informática jurídica*, data del año 1948, cuando el norteamericano Norbert Wiener publicó su libro *cybernetics* dentro del cual se expone el paralelo existente entre el cerebro humano y las máquinas electrónicas de cálculo, en su título le daba el nombre a una nueva ciencia que definió como la "ciencia del control y de la comunicación en los animales y en las máquinas"¹⁷. Es Wiener el primero en percibir el paralelo existente entre el cerebro humano y las máquinas electrónicas de cálculo¹⁸.

En 1949, un año después de la aparición del libro de Wiener, el juez norteamericano Lee Loevinger, publicó un artículo titulado "*Jurimetrics*"¹⁹, dentro del cual de forma primigenia se hace mención al concepto de *jurimetría*, término designado al conjunto de investigaciones lógica-matemáticas y estadísticas dirigido a los diversos tipos de

¹⁷ FROSINI, V. (1998). *Informática y Derecho*, Temis, Bogotá. p.41.

¹⁸ VARGAS OSORNO, T. G. (1993). *Manual de introducción a la informática jurídica*. Bogotá: Universidad Externado de Colombia. p. 38.

¹⁹ SÁNCHEZ MAZAS, M. (n.d.). *Panorama general de la Informática Jurídica*. Fundación para el Desarrollo de la Función Social de las Comunicaciones. FUNDESCO. p.7.

análisis de la información jurídica y su manejo informático de manera automatizada y racionalizada²⁰.

Esta disciplina surgió como una respuesta a las exigencias del sistema jurídico del *common law*, en el que el precedente obliga a los tribunales en la resolución de un caso concreto, se encontraban con un número de resoluciones abrumador y de carácter creciente; La posibilidad de buscar el precedente de una manera automática fue reveladora para los juristas norteamericanos²¹.

En cambio el carácter de los sistemas jurídicos de los países de Europa continental hace que los juristas sean poco sensibles al problema de la jurimetría en sentido estricto, la previsión del comportamiento futuro del juez. Pero incluso en los países anglosajones ha quedado probado la ineficacia de esa misión de la jurimetría: la administración de justicia es mucho más compleja que una regla matemática, y en los jueces influyen factores psicológicos reñidos con una aplicación rigurosa de precedentes por parte de la computadora, ha sido precisamente un juez norteamericano, Honfel, quien ha puesto de relieve que la formulación matemática que pretendía la jurimetría no era ni humana, ni jurídicamente válida²².

En su primitiva concepción, la Jurimetría como "medida del derecho" o posibilidad de establecer mediante un análisis matemático de datos estadísticos de los precedentes judiciales cuáles debían ser las soluciones a adoptarse por los tribunales,

²⁰ VARGAS OSORNO, T. G. op. cit. p. 39

²¹ LÓPEZ-MUÑOZ GOÑI, M. (1979). *Curso de informática jurídica*. Madrid: Fundesco. N. 8 de los cuadernos de Teleinformática Jurídica, Segunda Edición. p. 11-12.

²² HOFELD, W. (1969). *Concetti giuridici fondamentali*. Turin: Giulio Einaudi. p.63

fue abandonada frente al hecho decisivo de que sus previsiones resultaban muchas veces desmentidas por la realidad²³.

Otra postura jurimétrica que tampoco tuvo éxito es la que propugna por la utilización del computador, no para una función de decisión sino para una función de cálculo aritmético o matemático de normas. Se dice que el lenguaje es muy impreciso dando lugar a muchas confusiones e imprecisiones en la legislación, pudiendo el ordenador solucionarlas convirtiendo en signos y fórmulas matemáticas todos los conceptos jurídicos de las leyes²⁴.

Paralelamente con el tratamiento electrónico se había difundido en Europa la parte teórica de la Cibernética. La insatisfacción por los resultados concretos ofrecidos por la jurimetría y la presencia de instrumentos teóricos atractivos, como los ofrecidos por la cibernética teórica, hicieron que en Europa los estudios puramente empíricos de tipo loevingeriano se unieran con estudios de tipo puramente teórico, con el resultado que entre 1966 y 1969, con la denominación de "Cibernética y derecho" se designaran, tanto las encuestas de estadística judicial que recurrieran al ordenador, como los estudios de lógica formal aplicada al derecho, tanto los trabajos puramente computacionales que de alguna manera tuvieron que ver con normas jurídicas, como las investigaciones de filosofía del derecho que recurrieran a esquemas teóricos provenientes de la cibernética o de la teoría de la información²⁵.

²³ VARGAS OSORNO, T. G. op. cit. p. 40

²⁴ *Ibidem*.

²⁵ *Ibidem*.

La fusión interdisciplinar, característica de la cibernética en el sector jurídico que se transformó en confusión de disciplinas, esa falta de claridad contribuyó, a que los estudiosos perdieran interés, arrojando un cierto descrédito sobre la cibernética y el uso de la computadora²⁶

En 1968, Mario Losano con el interés de poner un poco de orden a esta situación, propuso sustituir el término *jurimetría* por el término *Iuscibernética*²⁷, que es la ciencia genérica dentro de la cual hoy en día se ubica la informática jurídica, por lo cual se ha logrado hacer la importante fusión de estas dos ciencias. De ahí, que propone la adopción del término *iuscibernética* para designar con carácter general a cualquier aplicación de la cibernética al derecho.

De acuerdo con lo manifestado por el profesor Losano, este vocablo no evoca la idea de cuantificación de los fenómenos jurídicos, como ocurre con la *jurimetría*, sino que se individualiza el método con el cual van a ser encarados los distintos aspectos del derecho. En ese sentido la *iuscibernética* constituye un conjunto de procedimientos técnicos para aproximarse al estudio del fenómeno jurídico. Ya no se trata de medir el derecho, sino de poner en relación el IUS con la Cibernética²⁸.

²⁶ Ibídem. p. 41.

²⁷ LOSANO, M. G. (1982). *Introducción a la Informática Jurídica*. España: Universidad de la Palma Mallorca. Facultad de Derecho.

²⁸ LOSANO, M. G. (1969). *Giuscibernética Macchine Cibernetici nel Diritto*. Torino: Einaudi. p. 21.

1.4. Clasificación

Una vez establecida la definición de *informática jurídica*, encontramos que a su vez esta se subdivide en tres ramas a saber: Informática jurídica documental, informática jurídica de gestión y control e informática jurídica metadocumental. A partir de esta división, es menester mencionar que nos dedicaremos al estudio de la informática jurídica documental, rama que se ocupa del estudio de los bancos de datos jurídicos y la organización del ordenamiento jurídico de un Estado determinado, tema objeto de esta tesis.

Siguiendo el hilo conductor, empezaremos con la definición de la Informática jurídica documental, que es entendida como la rama de la informática jurídica prístina “que se ocupa del tratamiento automatizado de los documentos jurídicos, especialmente los relacionados con legislación, jurisprudencia y doctrina²⁹.”

A partir de esta se construyen las bases de datos documentales sistemáticamente organizados, que almacenan y recuperan de manera automática grandes cantidades de información jurídica escrita y en la actualidad, por qué no decirlo, también información jurídica audiovisual.

Este primer sector es el más antiguo de toda la informática y por ende, de la informática jurídica, sus orígenes están en la presentación que hizo el profesor John

²⁹ VARGAS OSORNO, T. G. op. cit. p. 42.

F. Harty, de la Universidad de Pittsburg, en la American Bar Association en 1960³⁰, de un sistema que había creado en el año de 1959.

En el pasado las personas que se interesaron por esta nueva ciencia, centraron toda su atención, en los problemas que se presentaron en el tratamiento de la información y de los instrumentos lingüísticos para la solución de los mismos. Así mismo, gran preocupación se tuvo por las limitaciones de espacio en disco duro que hizo que los estudiosos del tema se acomodaran y estudiaran esta ciencia teniendo en cuenta todo este tipo de inconvenientes. Estudiaremos en su orden estos inconvenientes.

1.5. Sistemas De Tratamiento De La Información Jurídica

Es importante ahora, hacer referencia a los denominados sistemas de tratamiento de la información jurídica, que se constituyen como el medio que permite realizar la búsqueda de documentos con el establecimiento de criterios de información predefinidos que facilitan la organización y depuración de la información y la recuperación de la misma.

“Se entiende por sistemas de tratamiento de la información, aquellos de carácter informático utilizados para establecer los criterios de búsqueda de los documentos³¹”. En otras palabras, los sistemas de tratamiento de información jurídica conciernen a la metodología empleada para el tratamiento de la información

³⁰ VARGAS OSORNO, T. G. op. cit. p. 43

³¹ LÓPEZ MUÑOZ-GOÑI, M. (1984). *La informática jurídica documental*. Madrid: Díaz de Santos. p. 26

jurídica, ella proviene de ciertas técnicas desarrolladas para la búsqueda automática de la información « automatic retrieval »³².

Dentro de los sistemas definidos anteriormente, en los años sesenta en el ámbito internacional, hace su aparición los Sistemas de Recuperación de Información (SRI), que como su nombre lo indica, permite la recuperación de documentos de manera automática, por lo cual se pueden adaptar al ámbito jurídico.

Ahora bien, varios son los sistemas que se han ofrecido en el mundo para el tratamiento de la información jurídica y en especial para facilitar y organizar su uso, estos son: El sistema referencial, el sistema de full-text, el denominado abstract jurídico y el método de indexación. Es menester mencionar que estos sistemas se constituyen como el medio que permite realizar la búsqueda de documentos con el establecimiento de criterios de información predefinidos que facilitan la organización, depuración y recuperación de la misma.

1.5.1. Sistema referencial

Cualquiera que sea el contenido del banco de datos, el sistema limita su trabajo a proporcionar una simple referencia al documento, con la finalidad que el mismo pueda ser buscado en otro soporte. Es el sistema más económico y fácil de hacer para sostener y recuperar la información. Pese a su sencillez, cabe mencionar que este

³² SALAZAR CANO, E. (1978). Informática jurídica. Instrucción programada y documentación automática. Caracas: Universidad de Carabobo. p. 82.

sistema puede llegar a ser de gran utilidad como soporte complementario de uno más amplio y complejo³³.

1.5.2. Sistema full-text

El sistema *full-text* texto integral, es aquel que representa el documento original a partir del conjunto de palabras significativas contenidas en el mismo, de forma que la información incluida en la base de datos bien podría ser recuperada a través de todas las palabras relevantes contenidas en el texto, con excepción de las denominadas palabras nulas.³⁴

*En otras palabras podemos decir que el sistema de full-text (texto integral) es la entrada completa en la memoria del computador del documento jurídico, lo que trae como consecuencia la onerosidad, a causa de la gran cantidad de memoria utilizada, asunto que en el pasado constituía un límite en la creación de bancos de datos. Además, es muy factible obtener información distinta de la que es requerida. La localización de la información deseada en este sistema, se hace con base en palabras claves o palabras llaves existentes en el mismo texto*³⁵.

Como todos los sistemas de información jurídica, este presenta ventajas y desventajas, como ventajas, podemos mencionar que permite al usuario obtener no

³³ VARGAS OSORNO, T. G. op. cit. p. 44

³⁴ Son aquellas que por su naturaleza se considera que carecen de valor para efectos de recuperación y por tanto son omitidas, tanto de los documentos ingresados a un banco de información como de las consultas formuladas al sistema.

³⁵ VARGAS OSORNO, T. G. op. cit. p. 47.

sólo la indicación de la ubicación de los textos, sino el texto mismo³⁶. Igualmente, elimina la limitación de tener que usar un lenguaje cerrado para acceder a la información, de manera que se pueda interrogar mediante un lenguaje abierto o natural³⁷. Lo que admite prescindir de personal especializado en la materia.

Lo anterior, trae como consecuencia, la ausencia de trabajo manual de elaboración de documento, pues éste se indexa por sí mismo³⁸.

Como desventajas de este sistema, tenemos que al ser una transcripción, omite conceptos implícitos del documento fuente, por los cuales también se puede describir información y por lo tanto consultarse. Como consecuencia, se genera la posibilidad de que exista ruido informático, cuando los términos y el concepto del documento no van calificados³⁹ y como consecuencia de la imposibilidad de recuperación de información específica, no solo es posible que se produzca ruido, sino también silencio informático⁴⁰.

También, en este sistema se ocupa espacio innecesario en el disco duro del computador, debido a las inútiles repeticiones de términos que podrían reagruparse de manera lógica por medio de un análisis adecuado.⁴¹ Igualmente, se hace difícil y lenta la consulta para localizar los puntos de interés, ya que el dispositivo de salida del ordenador reproduce todo el texto⁴².

³⁶ *Ibíd.* p. 49

³⁷ Lenguaje abierto significa que no existe una lista previa, cerrada y obligatoria de descriptores que deba utilizar el analista para indexar el documento.

³⁸ VARGAS OSORNO, T. G. op. cit. p. 50.

³⁹ *Ibíd.*

⁴⁰ *Ibíd.*

⁴¹ *Ibíd.* p. 51.

⁴² *Ibíd.*

Varios autores y estudiosos de la informática Jurídica, como Edgar Salazar Cano, Alberto Serrano Pirella, Ana Manrique de Carrión y Alberto Serrano entre otros⁴³, coinciden en considerar al sistema de full text, ideal para el tratamiento de información legislativa.

1.5.3. Abstract jurídico

Como un tercer sistema, encontramos el *abstract jurídico*, otra alternativa para el tratamiento de la información jurídica, el *abstract* es definido desde un punto de vista jurídico informático, como el conjunto de descriptores que definen el contenido de un documento y permiten su recuperación⁴⁴.

En el mismo sentido el *abstract* es definido como un resumen o sinopsis de los contenidos de un documento que conserva la estructura básica de la información contenida en el mismo, este sistema se presenta como aquel conjunto de descriptores que definen el contenido de un documento y permiten su recuperación, conservando la estructura básica de la información contenida en el mismo.

Las ventajas que ofrece este sistema, residen en permitir la consulta de un banco de datos en un grado general o específico de acuerdo con lo querido por el usuario. En el cual se utiliza un lenguaje abierto y libre. Además de lo anterior, al señalar los descriptores que conforman el *abstract*, se discrimina la información que no tiene importancia, relevancia jurídica o que es repetitiva. Y finalmente, admite que la

⁴³ Ibídem. p. 49.

⁴⁴ Ibídem. p. 52.

información referida con un mismo concepto, órgano, autoridad o institución que se encuentra desarrollada en varios artículos y capítulos, se reagrupe.

En los problemas de técnica legislativa, combinados con el uso de instrumentos lingüísticos: Léxico y tesauros jurídicos, el *abstract* resulta exitoso en la solución de los mismos; especialmente, cuando hay reglamentación enorme de instituciones jurídicas que hace que se presente duplicidad de disposiciones. En cuanto a las aplicaciones de abstract, el programa de automatización UNAM-IURE, es pionero en su aplicación para el tratamiento de la información documental legislativa, siendo el IRETIJ de Montpellier, Francia, quien lo introdujo y utilizó con muy buenos resultados en materia de Jurisprudencia⁴⁵.

1.5.4. Método de indexación

Para finalizar, el método de indexación o *keyWord*, “consiste en asignar a un conjunto de información y algunos de sus subconjuntos, ciertos descriptores de tipo general (tomados de una lista rígida) a los que se debe encuadrar toda la información analizada, con la finalidad de contar con puntos de contacto homogéneos de la información, según su especie, teniendo en cuenta un enfoque determinado⁴⁶”.

Indexar apunta a calificar la información por conducto de la elaboración de fichas de análisis sobre lo contenido en el documento fuente, asignando el descriptor

⁴⁵ Ibídem. p. 53.

⁴⁶ CÁCERES NIETO, E. (1984). *Teoría del abstract legislativo*. México: Universidad Nacional Autónoma de México. UNAM.

apropiado para el caso en concreto. Generalmente el descriptor es seguido de un resumen que enuncia someramente el contenido del soporte documentario⁴⁷.

Siempre que hablamos de indexación, nos referimos a un sistema de lenguaje cerrado, los cuales se basan en una lista de descriptores que son empleados tanto para tener acceso a la información contenida en la fichas de análisis, como para la elaboración de éstas⁴⁸.

Los descriptores juegan un papel fundamental, no son puestos de manera arbitraria por el analista de información, sino que son tomados de un tesoro jurídico, a partir del cual se realiza la calificación de la información.

En este punto es necesario mencionar que los *descriptores* o *keywords*, juegan un papel trascendental dentro de este método primigenio de la informática jurídica, que dicho sea de paso y como es bien sabido, persiste aún dentro de la informática jurídica de la actualidad.

Un *descriptor* es una palabra clave que tiene la capacidad de definir la totalidad o una parte del contenido de un documento con el fin de discriminar y organizar la información de forma que facilite su posterior ubicación bien sea en un archivo manual o en uno informatizado, que se emplea con el fin de tener acceso a la información contenida en una base de datos, de manera que mal se haría en otorgar denominaciones arbitrarias al contenido de la información, motivo por el cual los descriptores son tomados de los *tesoros jurídicos*.

⁴⁷ VARGAS OSORNO, T. G. op. cit. p. 45

⁴⁸ CÁCERES NIETO, E. op. cit.

De otra parte como desventaja, se presenta el hecho de estar condicionado al uso de un sistema cerrado de recuperación de información, es decir, que para poder obtenerla se debe interrogar por alguno de los descriptores que se encuentran en el *thesaurus*⁴⁹; esto conduce a la exigencia de un cuidadoso análisis humano, requiriendo para ello personal especializado que designe los descriptores correctos a cada documento fuente⁵⁰. Por lo anterior, resulta importante la labor que desempeña el analizador toda vez, que un documento mal indexado se tornará irrecuperable, de forma que se puede llegar a incurrir en el denominado *ruido informático*, que trae como consecuencia información no requerida y la imposibilidad de recuperar toda la información sobre un punto específico, esto último es conocido como *silencio informático*.

Una ventaja de carácter financiero se dilucida en este sistema de indexación, por cuanto “reduce el material de base y por consiguiente el costo de alimentación⁵¹”.

1.6. Instrumentos lingüísticos

Las manifestaciones del lenguaje humano constituyen la denominada lingüística. El lenguaje se caracteriza por ser multívoco e impreciso; entonces el primer y más importante aspecto del lenguaje es su no exactitud, su indeterminación, lo que significa que el lenguaje como tal no contiene todos los elementos que permitan realizar una comunicación completa, unívoca y no sujeta a interpretaciones. El

⁴⁹ VARGAS OSORNO, T. G. op. cit. p. 46.

⁵⁰ *Ibidem*.

⁵¹ LE BOULENGUE DE GALARD, M. (n.d.). *El tratamiento electrónico de la documentación jurídica*. Anuario de la Facultad de Derecho. p. 189-190

significado depende del contexto o campo lingüístico, así como de la situación dentro de la cual se desarrolla la comunicación.

En este punto, se hace necesario mencionar que, como una solución a estos problemas que se derivan del lenguaje surgen los instrumentos lingüísticos, que se definen como aquellos que tienen por objeto resolver los problemas gramaticales, con el fin de lograr la recuperación de información por ideas independientemente de la forma en que se expresen.

El desarrollo de un sistema de información exige un importante trabajo sobre el vocabulario. Son instrumentos lingüísticos: La Fonología (sonidos del lenguaje), Sintaxis (estructura del lenguaje), Semántica (estudio del significado). Algunos de los problemas que se presentan en un sistema documental en relación con el lenguaje son: La polisemia, homonimia, sinonimia, la analogía, antinomia, palabras vacías, formas sintácticas, vaguedad y ambigüedad. En el desarrollo de la informática jurídica documental, para solucionar estos problemas que presenta el lenguaje se han desarrollado dos instrumentos lingüísticos: El tesauro y el léxico. De los cuales nos ocuparemos a continuación.

1.6.1. Tesauros

Existe diferencia entre el lenguaje natural y el lenguaje que las máquinas comprenden, es decir, el ordenador es una máquina que no goza de capacidad para entender el sentido de un texto, ya que el solo contrasta dentro de sus archivos la existencia de caracteres sin lograr hacer razonamiento alguno, finalmente es una

máquina. Los datos jurídicos son fundamentalmente lingüísticos y por lo mismo es necesario crear una interface entre el lenguaje que la máquina logra comprender y el lenguaje humano. Uno de estos instrumentos lingüísticos de gran importancia y que cumple un papel fundamental en la solución de estos problemas es el tesoro, vocablo que proviene del griego y significa acumulación, almacenamiento o tesoro.

“J.Mesiger fue el primero en utilizar la palabra TESAUROS, en un sentido específico, como almacén de conocimientos aplicado a un diccionario estructurado por la correlación de lenguas, en su obra thesauruspolyglotys. El concepto de thesaurus como correlación de las palabras se incorpora en el año de 1736, en Inglaterra, al lenguaje común, expandiéndose al resto del mundo como nuevo vocablo, con sentido propio”⁵².

Dentro de las diferentes definiciones de tesauros, encontramos la que lo define como un “conjunto de palabras y o expresiones descriptoras que tienen la posibilidad de estar aclaradas en su alcance, relacionadas entre sí y aglutinadas bajo un criterio de ordenación que suele ser el alfabético”⁵³.

Por su parte, H. BaverBernet, define a los tesauros como “un léxico jerarquizado que compone una red de relaciones lingüístico-jurídicas, con un conjunto estructurado de descriptores⁵⁴”. En igual sentido, Hector Fix Fierro, define el tesoro

⁵² LÓPEZ MUÑOZ-GOÑI, M. (1984) op. cit.

⁵³ SUÑÉ LLINÁS, E. (1988). *Los tesauros en bases de datos de usuario*. En la obra colectiva: Actas del congreso de informática y derecho. Roma. p. 1-27

⁵⁴ BAUER-BERNET, H. (1982). *Legal thesaurus and data processing*. Amsterdam.

como “un léxico jerarquizado que comprende una red de intercomunicaciones, de exclusiones, de discriminaciones y de proximidades semánticas”⁵⁵.

Igualmente, se ha definido al tesoro como un instrumento de clasificación y ordenación de los textos y documentos en función de sus relaciones semánticas de analogía, vecindad, inclusión o jerarquía. El tesoro permite orientar las búsquedas de los usuarios facilitándoles la información deseada; su eficacia se mide por la capacidad para desarrollar esa función con la mayor dosis posible de exhaustividad y precisión⁵⁶.

Podemos concluir, que el tesoro es un instrumento lingüístico, que aclara las relaciones semánticas no sólo entre las palabras sino también a nivel de los conceptos e ideas que lo integran y permite conocer las relaciones de jerarquía, inclusión, preferencia y asociación de los conceptos. Los tesauros sirven de medio de ayuda al usuario para resolver los problemas de analogía, antinomia y el aislamiento semántico de las palabras polisémicas, al momento de la interrogación dentro de un banco de datos.

Que exista un vocabulario organizado y estructurado, que las palabras que lo integran estén correlacionadas por diferentes vías semánticas o lingüísticas, sintagmáticas, paradigmáticas y asociativas y que se refiera a una disciplina en particular, son requisitos *sine quanon* para que un tesoro sea considerado como tal.

⁵⁵ FIX FERRO, H. (1984) op. cit. p. 87.

⁵⁶ PÉREZ LUÑO, A. E. (1996). op. cit. p. 152.

Tradicionalmente los tesauros han sido utilizados en la catalogación bibliográfica, sin embargo, actualmente es un instrumento de gran utilidad en el campo de la información documental⁵⁷.

1.6.1.1. Clasificación de los tesauros

La clasificación de los tesauros, puede ser de diversos tipos

1.6.1.1.1. Según su utilización:

- *Tesauros manuales*: Aquellos que se presentan en forma impresa para su utilización manual⁵⁸.
- *Tesauros automatizados*: Son aquellos que se encuentran incorporados a un computador y pueden ser:
 - ✓ Tesauros para la información. Son aquellos que se incorporan a una base de datos especial, independiente de la principal, para que el usuario pueda consultar cuál es el vocabulario utilizado en la base que va a interrogar.
 - ✓ Tesauros para la recuperación: Son aquellos que permiten el acceso a la base de datos a través del mismo tesoro, quien estará

⁵⁷ VARGAS OSORNO, T. G. op. cit. p. 62.

⁵⁸ LÓPEZ MUÑOZ-GOÑI, M. (1984) op. cit.

interrelacionado con la base principal. En consecuencia es factible interrogar por medio del thesaurus⁵⁹.

En este punto se hace necesario resaltar que existen dos sistemas para la recuperación de la información documental: Un primer sistema es el de *palabra clave dentro del texto*, conocido con el nombre de SISTEMA KWIC (Key Word in context), en el cual el computador recupera la información buscando las palabras que se le dan como argumento de búsqueda en el texto de los documentos que almacena en sus archivos. El segundo es el sistema KWOC (key Word outof context), que busca la información requerida a partir de palabras que se colocan fuera del texto por quien recopila la información, las que corresponden al contenido del mismo⁶⁰.

En cuanto al primer sistema, encontramos que con el fin de recuperar la información a partir del texto del documento archivado, se pueden presentar modificaciones para facilitar su localización. Surgen a partir de allí, tres tipos de tesauros: los elaborados para buscar la información partiendo de las palabras naturales en que está redactado el documento, aquellos que permiten la búsqueda a partir de un lenguaje predefinido en el que se transcribe el documento o se hace un resumen del mismo, y finalmente, aquellos tesauros contruidos a partir de la búsqueda con palabras asignadas por fuera del texto del documento almacenado, que corresponden a los temas tratados⁶¹.

⁵⁹ Ibidem.

⁶⁰ GIRALDO ÁNGEL, J. (1990). *Informática jurídica documental: manual para el sistema universitario de informática jurídica*, SUIJ. Temis. Bogotá. p. 10

⁶¹ Ibidem, p. 11

Dentro del primer tipo de tesauro, es decir, el de *recuperación de información a partir de las palabras en que está redactado un documento*, surge el denominado listado de palabras naturales, que es utilizado por algunos sistemas de informática documental⁶², en los cuales se crea un listado de palabras que reposan en los diferentes textos almacenados sobre los cuales se pretende su ubicación.

Cuando se emprende la búsqueda por medio de este sistema de recuperación de información, basta señalar la palabra para que automáticamente el computador arroje todos los documentos que contiene el documento.

Con el fin de impedir la pérdida de documentos, cuando dentro de la búsqueda no se encuentran explícitamente la palabra utilizada, bien sea porque la palabra tiene sinonimia, o pueda ser susceptible de traducción a otras palabras sustituyéndolas o combinándolas con otras, surge como medio de solución el denominado *tesauro semántico*, que principalmente busca la convergencia de los términos del lenguaje, a estos, se les ha otorgado la denominación de “semas”, que son definidos como pequeñas unidades de significación producto del análisis. Verbigracia: De cujus, causante, finado, occiso, muerto, difunto, fallecido, extinto, ajusticiado, cadáver.

En cuanto a las condiciones de selección, los investigadores de la oficina del Maximario de la Corte Suprema de Casación de Italia, han determinado ciertas medidas para establecer los términos que constituyen semas, a saber son: Que tengan un significado elemental, que su significado sea lo más distinto posible de todas las

⁶² Ibídem, p. 12.

otras “semas”, que pueda traducir un número apreciable de otras palabras, sustituyéndolas o entrando en la combinación con otros “semas”⁶³.

Finalmente, es importante resaltar que este tesoro modifica la estructura del sistema tradicional de búsqueda y proporciona al usuario del mismo, la oportunidad de obtener información con variaciones morfológicas del lenguaje de los términos asociados con la palabra investigada.

Otro de los tesauros para la recuperación de información, que surge a partir de las palabras en que está redactado el documento, es el *gramatical*, este contiene la totalidad de concordancias sobre la palabra conforme a sus diferentes géneros, números, tiempos, conjugación, variaciones de personas, sustantivos, adjetivos, verbos irregulares y diferentes modos en los verbos. El tesoro gramatical, otorga la posibilidad de recuperar información introduciendo cualquier argumento de búsqueda, para que el computador arroje como resultado los documentos que contienen la palabra y sus diferentes desinencias.⁶⁴

En cuanto al segundo tipo de tesauros, es decir, aquellos que permiten la *búsqueda a partir de un lenguaje predefinido en el que se transcribe el documento o se hace un resumen del mismo*, con el fin de solucionar los problemas que suscita la recuperación de información partiendo del lenguaje natural contenido en el documento, lo que se realiza aquí, es la incorporación sintética del mismo,

⁶³ Ibídem, p. 13.

⁶⁴ La desinencia es un morfema dependiente que se añade a la raíz de una palabra en las lenguas sintéticas para indicar diversos aspectos gramaticales, como el género, el número, la persona, el caso, el tiempo o el modo.

eliminando los caracteres innecesarios, de forma que el listado de palabras será mucho más técnico y reducido, pero ligado al lenguaje en que fue escrito⁶⁵.

Ahora bien, dentro de este segundo tipo, se encuentran aquellos tesauros elaborados a partir de un lenguaje predefinido, que al igual que el listado de palabras construido a partir de resúmenes, busca solucionar los problemas que presenta la recuperación de información por medio del lenguaje natural, estos cuentan con una estructura mucho más rígida que resume el documento en párrafos separados. Los más reconocidos son el Iretij y el Unam-jure.⁶⁶

Finalmente, encontramos los tesauros contruidos a partir de la búsqueda con palabras asignadas por fuera del texto del documento almacenado, que corresponden a los temas tratados. Aquí se presenta una recuperación de los temas a que hace relación el documento, contrario a los otros tesauros que ofrecen una búsqueda o recuperación de información mediante las palabras utilizadas para presentarlos.

⁶⁵ GIRALDO ÁNGEL, J. (1990). Op.cit., p. 14.

⁶⁶ Ibídem. "El Iretij es un banco creado por el "Institute de Recherches et d'Études pour le Traitement de l'Information Jurudique", siguiendo la metodología diseñada por el profesor PIERRE CATALA para el manejo de la jurisprudencia. La unidad informática es el "abstract", que es un resumen, en párrafos separados, de cada uno de los problemas que se resuelven en la providencia que se va a incorporar al archivo del computador, redactados en un "léxico" predefinido que tiene 10.000 palabras. (...)

En el Unam-jure la unidad informática es la ley, el decreto, etc., redactando un párrafo por cada uno de los temas que se ocupe el estatuto jurídico objeto de análisis.

Los párrafos tienen una estructura de forma rígida, pues se componen de frases, cada una de las cuales juega un papel definido: La primera contiene el nombre de la ley, decreto, etc., especificando los artículos que se refieren al correspondiente tema; la segunda indica el tema del párrafo; las siguientes enuncian los desarrollos que la ley o decreto dan al tema".

En otros términos, este último tesoro ofrece una recuperación de información a partir del contenido conceptual de los documentos y no de las palabras en que está escrito⁶⁷.

Esto a su vez permite solucionar en gran medida el ruido informático y el silencio informático. No obstante, para tales efectos exigirá un análisis de la información de forma exhaustiva que bien puede llegar a resultar dispendiosa y producir distorsión en la determinación del documento y en la asignación de sus descriptores, por lo cual en el uso de este tipo de tesauros se requiere en mayor medida el conocimiento jurídico de la persona que realiza el análisis de la información.

1.6.1.1.2. Según su idioma

- *Tesauros monolingue*: Es usado en una sola lengua⁶⁸.
- *Tesauros bilingue*: Es aquel que establece relaciones no sólo entre los diferentes descriptores, sino también con su correspondiente traducción en dos o más idiomas⁶⁹.

En el campo del derecho, para GuyMazet, se presenta una confrontación de los órdenes jurídicos en cuestión, además de la consecuente ruptura en la comunicación por el problema de las diferentes lenguas⁷⁰.

⁶⁷ Ibídem. p. 16.

⁶⁸ VARGAS OSORNO. T. G. (1990). Op.cit. p. 65.

⁶⁹ Ibídem.

⁷⁰ MAZET, G. (1983). *Acces en differentes langages a un Corpus Documentaire Problematique et methodology*. In III Congresso Internazionale su L'informática Giuridica e la Comunita Nazionale ed Internazionali. Roma.

1.6.1.1.3. Según su estructura

- *Tesauros lineales*: “Son los que presentan relación de palabras sin conexiones de mayor grado. Pueden considerarse como simples vocabularios. A lo sumo tienen relaciones de género-especie pero sin interrelacionarse unos con otros”⁷¹.
- *Tesauros arborescentes*: “Son aquellos que se encuentran estructurados en forma ramificada, de manera que por las relaciones ascendentes o descendentes que se plantea un descriptor puede tener un padre y varios hijos”⁷².
- *Tesauros rediculares*: “Son aquellos en los que la información está entrecruzada, de tal manera que los descriptores pueden tener varios padres y varios hijos, relacionándose siempre por vía ascendente o descendente, pero de forma indefinida”⁷³.

De esta manera la información puede irse obteniendo paso a paso, hacia arriba o hacia abajo, según interese. Al encontrarse, todos los descriptores relacionados, este proporciona la información de manera más completa.

⁷¹ VARGAS OSORNO. T. G. (1990). Op.cit., p. 66.

⁷² *Ibídem*.

⁷³ *Ibídem*.

1.6.1.1.4. Según su contenido

- *Tesaurus monodisciplinares*: “Son aquellos que sólo se ocupan de una disciplina”⁷⁴.
- *Tesaurus interdisciplinares*: “Son aquellos que intercomunican a bases de datos de diversas disciplinas”⁷⁵.
- *Tesaurus temáticos*: son aquellos en los que se recupera la información a partir del contenido temático de los documentos. Esta clase de tesauros, es utilizada para la elaboración de los bancos de datos jurídicos de la Universidad Externado de Colombia. Los consideramos de especial importancia para el presente estudio, dado que es la base fundamental para la creación de ontologías aplicadas al derecho; tema que se analizará con profundidad en la tercera parte de esta tesis.

En los tesauros que se elaboran para recuperar información a partir de su contenido temático, resulta ineludible mencionar sus características principales, toda vez que gozan de unas cualidades excepcionales que los diferencian de los demás tesauros.

Son características de los tesauros temáticos las siguientes:

⁷⁴ *Ibíd.*

⁷⁵ *Ibíd.* p. 67

1. Cuando un tesoro se elabora con el fin de reflejar el contenido conceptual del documento jurídico, surge la necesidad de incluir los términos con los cuales se permita acceder a la información, estos términos son los llamados “*descriptores*” y cuentan con la particularidad de ser palabras fuera del texto mismo, que buscan expresar su contenido. Por esta razón, la primera característica con que cuenta un tesoro temático es que se encuentra construido a partir de *descriptores*, los que podemos definir como aquella palabra que tiene contenido conceptual autónomo y representa un concepto jurídico.

Jaime Giraldo Angel, define el término descriptor como, “toda palabra que tenga contenido conceptual propio, aunque comparta con otros descriptores elementos comunes. Son descriptores, por ejemplo, términos como derecho civil, contrato, arrendamiento, compraventa, arrendatario, canon, por cuanto todos ellos tienen un contenido conceptual autónomo, así algunos de ellos pertenezcan a un mismo género próximo y así algunos de ellos tengan un contenido muy amplio y otros muy restringido.”⁷⁶

Los descriptores se caracterizan por estar compuestos por una o varias palabras, cuando el descriptor está formado por un conjunto de palabras que constituyen una unidad conceptual se le denomina *sintagma*, son ejemplos de sintagmas: Lesión enorme, derechos humanos, compraventa bienes inmuebles, libertades públicas, etc.

En esta instancia, se hace necesario diferenciar el sintagma de la unión de varios descriptores, por cuanto en el primer caso, esto es, el sintagma, el significado surge

⁷⁶ GIRALDO ÁNGEL, J. (1990). Op.cit., p. 24.

de valorar sus palabras compuestas, ya no individualmente, como sucede en el caso de la unión de varios descriptores donde cada uno de estos términos tiene una unidad conceptual. Un ejemplo de la unión de varios descriptores podría ser “contrato” y “cesión”.

Los descriptores a su vez, debido a la gran cantidad de información que pueden llegar a contener, pueden ir acompañados por un restrictor, que son aquellos términos que permiten, (como su misma palabra lo dice) restringir el contenido de la información, de tal forma que la limite mediante unos criterios de búsqueda específicos, un ejemplo práctico dentro de un tesoro jurídico lo podemos ver cuando se introduce en la búsqueda el descriptor “contratos”, que arrojaría resultados enormes, por cuanto traería todas las clases de contratos que existen, por lo cual, con el fin de delimitar el estudio y particularizarlo, se puede incluir los denominados restrictores que genera como resultado una información restringida, que permite traer exactamente lo que el usuario requiere, sin que se cree ruido informático; así pues, al ingresar en la búsqueda el descriptor “contratos” acompañado del restrictor “requisitos” obtendrá la mencionada depuración de la información, ya que el sistema de información, solo traerá lo que se refiere a los requisitos de los contratos.

Ahora bien, los restrictores, no cuentan con autonomía conceptual, a diferencia de la unión de descriptores, con los cuales bien se podría realizar una búsqueda independiente, igualmente, su interpretación jurídica bien podría variar dependiendo del término al cual se refiere, verbigracia: “formas de terminación”, “agravantes”, “causales”, “naturaleza jurídica”, etc. Esto resulta relevante, a la hora de determinar que se puede entender por descriptor y por restrictor, mal se haría entonces en tener

como descriptor por ejemplo “naturaleza jurídica de la compraventa”, aquí lo correcto sería COMPRAVENTA – Naturaleza jurídica.

Todo descriptor, por referirse a un concepto, se enuncia con un sustantivo, “el cual se debe expresar en género masculino y número singular⁷⁷”, excepto cuando la misma naturaleza del término imponga el femenino, o cuando el término, se encuentre conformado por más de una palabra, se debe escribir en plural. Ejemplo: Libertades públicas.

2. Los términos que integran los tesauros temáticos que se utilicen para la consulta de un banco de datos jurídico, deben representar conceptos de la misma naturaleza y por lo mismo, todos los descriptores deben ser de contenido jurídico.
3. Los tesauros temáticos se presentan de tres formas: *Alfabética, sistemática y gráfica*. En la *estructura alfabética*, los descriptores aparecen organizados, como su nombre lo dice, en su orden alfabético y cada uno de ellos deberá establecer relaciones de jerarquía, preferencia y asociación, esto con el fin de catalogar y organizar la base de datos. Y además, cada descriptor debe aparecer con su correspondencia lógica con los demás términos del tesoro y estas relaciones pueden ser de especie a género; de particular a los general, etc.

En las relaciones de jerarquía, se incluye tanto lo que concierne al término superior al que pertenece el descriptor como las divisiones de este. El término superior está identificado con las iniciales T.G. que significa término general. Y a su vez las

⁷⁷ GIRALDO ÁNGEL, J. (1990). Op.cit., p. 25.

divisiones del mismo, se identifican con las iniciales T.E. que quieren significar término específico.

Dentro de las relaciones de preferencia, se escoge un término dentro de todos los que existan, teniendo en cuenta que bien se podría presentar la sinonimia para efectos de expresar un mismo concepto. Esto no obsta, para que los demás términos se encuentren referenciados, tanto sinónimos como antónimos; de manera que el usuario del Banco de datos puede buscar la información con el término conocido por él, y de forma automática, el sistema lo debe llevar al descriptor elegido y esto aparecerá con un VEASE.

Además de las relaciones de jerarquía y preferencia, existen las relaciones de asociación. En las cuales se sugiere la consulta de otros temas que complementan el tema objeto de estudio, Estos descriptores irán acompañados de las letras V.A. (Véase Además).

Ahora nos corresponde hablar sobre *la estructura sistemática*, como una de las formas de presentación del tesoro temático. Esta se encuentra compuesta por dos partes a saber: “en la primera los términos se ordenan en categorías o jerarquías, teniendo en cuenta los significados o interrelaciones lógicas, en la segunda los términos aparecen alfabéticamente en un índice que dirige al usuario a las partes apropiadas de la sección sistemática⁷⁸”.

⁷⁸ Ibídem. p. 26.

La última forma en que se presenta un tesauro temático es la *estructura gráfica*, aquí “los términos se pueden presentar de dos formas: En una estructura arbórea que establece la división de cada uno de los descriptores básicos en sus correspondientes especies o partes; o en una matriz, en la que el descriptor básico aparece en una cuadrícula central, de la cual se desprenden flechas que indican sus relaciones con los respectivos términos genéricos y divisiones específicas.⁷⁹”.

4. La cuarta característica de un tesauro temático, consiste en que *su formulación debe ser de carácter general y abstracto*, en especial si se trata de un banco de datos jurídicos de tendencia continental, en el cual las normas son positivizadas objetivamente, esta situación varía si nos encontramos frente a un sistema jurídico de corte anglosajón, dado que allí si cobra importancia la casuística y el precedente.

1.6.2. Léxico

Podemos definir al Léxico en sentido amplio y en sentido restringido⁸⁰.. En sentido amplio el léxico es el conjunto de signos lingüísticos contenidos en cada uno de los documentos del banco de datos, que en su totalidad conforman la base documental. En sentido restringido el léxico es el instrumento lingüístico de tratamiento del conjunto de signos lingüísticos contenidos en cada uno de los documentos del banco, organizándolos y dándoles cierta estructura, con el objeto de resolver los

⁷⁹ *Ibíd.* p. 27

⁸⁰ CASTAÑEDA PONCE, D., HERNÁNDEZ MARTÍNEZ, M., & GONZÁLEZ GALVÁN, J. (1989). *El léxico como instrumento lingüístico del sistema UNAM-JURE*. In J. Bilon, H. Fix Fierro, E. Cáceres , & S. Matute, *Diálogo sobre la informática jurídica*. México: Universidad Nacional Autónoma de México. UNAM. p. 245.

problemas lingüísticos que se presentan, para que el usuario pueda comunicarse eficazmente con la máquina al recuperar la información pertinente a su interrogación

Phillipe Nectoux y Henry Manzanares, definen al léxico como “aquel que comprende el conjunto de las palabras, es decir, de la cadena de caracteres significantes que el sistema reconoce”⁸¹.

Enrique Cáceres, lo define diciendo que consiste en la organización de todas las palabras con contenido informativo almacenadas en la computadora, en nociones y subnociones, bajo un criterio morfológico-semántico, con la finalidad de resolver los problemas de la sinonimia y formas sintácticas y ayudar a resolver el de la homografía o polisemia⁸².

Con el ánimo de comprender cómo por conducto del léxico se pretende resolver los problemas lingüísticos de sintaxis, sinonimia y algunos aspectos de polisemia, es necesario analizar, antes de explicar la organización del léxico, algunos conceptos:

NOCION: Es la agrupación de las palabras que tienen la misma raíz y que comparten un contenido semántico equivalente, con la finalidad de resolver en forma directa el problema de las formas sintácticas.

SUBNOCION: Consiste en la organización dentro de una noción⁸³ de las palabras con las que indistintamente se puede hacer alusión a la misma idea, con la doble

⁸¹ MANZANARES, H., & NECTOUX, P. (1988). *La informática al servicio del jurista*. Bogotá: Legis. p. 80.

⁸² CÁCERES NIETO, E. (1984). op.cit.

⁸³ En los sistemas de Iretij y del Unam-lure, una noción admite hasta diez (10) subnociones.

finalidad de que la computadora las trate de manera idéntica, para efectos de recuperación⁸⁴.

El léxico va a organizar a las palabras, primero las más significativas de la lengua, en unidades semánticas o nociones que constan de subnociones y por medio de las cuales el usuario puede recuperar la información⁸⁵. Este tipo de palabras son los verbos, los sustantivos, adjetivos y algunos adverbios.

Las preposiciones, conjunciones, artículos y algunos adverbios, (cuyo peso semántico es menor por ser considerados meros vínculos entre las palabras de una oración) se tratan de un archivo especial, donde de cierta forma se les aísla de las palabras activas, y a través de ellas no es posible recuperar la información, ya que se consideran palabras pasivas.

Los adjetivos numerales, a diferencia de las palabras pasivas sí pueden ser objeto de una interrogación del usuario.

En razón de la estructura morfológica de las palabras, cuando estas tienen la misma raíz presentan significados iguales o semejantes. Por ello, el léxico se estructura en nociones. La organización de la información documental se realiza con base en las palabras activas.

⁸⁴ CÁCERES NIETO, E. (1984). op.cit.

⁸⁵ Se les llama palabras activas, por ser signos lingüísticos a través de los cuales se puede recuperar la información contenida en los documentos.

La organización del léxico, basada en la estructura de las nociones, funciona permitiendo dos tipos de relaciones de las palabras: la automática (relación de equivalencia) y la opcional (relación de extensión).

El criterio básico con el que se constituyen las nociones es de tipo semántico. En caso de que dos o más palabras, aun siendo de distinta raíz, signifiquen lo mismo se incluirán dentro de la noción y en la subnoción que corresponda, como dos palabras que tengan la misma raíz y no compartan significado alguno, se les aislará en nociones distintas. A esto se le denomina Ruptura semántica total.

Son ejemplos de nociones con sus subnociones los siguientes:

- ✓ Contrato
- ✓ Contratos
- ✓ Contraten
- ✓ Contratados
- ✓ Contratar
- ✓ Contratado
- ✓ Contratarlo
- ✓ Contratación
- ✓ Contrataciones
- ✓ Contratista
- ✓ Contratistas

En cuanto al léxico y los problemas lingüísticos podemos mencionar:

Cambios verbales: Los problemas que se presentan con ocasión de los cambios verbales, el léxico los resuelve agrupando en una misma subnoción todas las formas verbales de la unidad morfológica semántico correspondiente.

Ejemplo:

Creación Creaciones Creó Creará Creado

El usuario podrá hacer su interrogación por cualquiera de las formas verbales y recuperar el documento, ya que el léxico relaciona de manera automática todas las palabras agrupadas en una misma subnoción.

Sinónimos: El léxico agrupa en la misma subnoción a las palabras con el mismo significado, pudiendo el usuario preguntar por cualquiera de las palabras sinónimas para recuperar el documento.

Ejemplo:

Panteón Panteones Cementerio Cementerios

Aquí se rompe con la regla general de agrupar en una misma subnoción las palabras con la misma raíz. Por ser sinónimos perfectos se tratan como equivalentes.

A diferencia de los análogos que tienen que separarse en nociones distintas, a pesar de compartir un significado similar.

Ejemplo:

Automóvil Coche Carro

Su recuperación solo es factible por medio de un thesaurus, por permitir éste relacionar palabras de distintas nociones.

Polisemia: Se plantea el problema de dónde ubicar a la palabra o palabras susceptibles de contenerse en dos o más nociones dependiendo de la acepción que presenten. Este problema el léxico lo resuelve separando a la palabra homógrafa en una noción aparte siempre y cuando en todas las palabras almacenadas en la computadora exista la otra u otras en que se pudiera ubicar.

Antonimia: Frente a los problemas presentados por la antonimia, el léxico los resuelve uniendo y separando las palabras según sea la clase de antonimia. Si es una antonimia morfológica se unirán a las palabras en una subnoción de la familia o unidad semántica correspondiente.

Ejemplo:

Disciplina = Indisciplina

La noción quedará:

- ✓ Disciplina Disciplinas
- ✓ Disciplinarse Disciplinar
- ✓ Interdisciplinario Interdisciplinarias
- ✓ Multidisciplinario Multidisciplinaria
- ✓ Indisciplina

Cuando es una antonimia general se separan las palabras en nociones distintas, por no pertenecer a la misma familia morfológica y no compartir significados similares.

Ejemplo:

Hombre DIFERENTE Mujer

1.7 . Recuperación de la Información

El ingreso de la información dentro de un banco de datos jurídico, busca que su contenido pueda ser utilizado. Con este fin, se hace necesario que la persona que lo consulte tenga conocimiento sobre las reglas lógicas y sintácticas para lograr encontrar la información pretendida.

Los primeros sistemas de interrogación de bancos de datos jurídicos fueron los sistemas Batch, es decir, las búsquedas en los ficheros que permitían una vez indicada la palabra, ir a buscar la dirección donde estaba almacenada, en todos los documentos del archivo, comparando simplemente los ficheros. De los sistemas Batch se pasó a los sistemas On line, con la posibilidad de

establecer interrelación a distancia a fin de precisar, con un diálogo entre usuarios y máquina la pregunta. Hoy los sistemas de interrogación son suficientemente sofisticados para ayudar al interrogador⁸⁶.

Para la comprensión del sistema lógico de tratamiento de la información es necesario incursionar en los temas de la *lógica formal, de los operadores lógicos y sintácticos*⁸⁷, sobre los cuales hablaremos someramente, siendo necesario tener claridad conceptual, con el fin de comprender cómo se recupera la información dentro de un sistema.

1.7.1. Lógica formal

Los computadores no están en capacidad de utilizar el lenguaje ordinario común; por el contrario, utilizan un lenguaje algorítmico de alta precisión y carácter unívoco. El fundador de la lógica fue Aristóteles, con la definición de los silogismos que ha perdurado casi sin modificación hasta el siglo XVIII.

Con Kant aparece el término Lógica formal, contrapuesto al de lógica trascendental. Para efectos de la aplicación de los computadores al derecho nos importa la lógica formal, la que tiene como objetivo fundamental introducir dentro del lenguaje ordinario un mayor rigor y precisión, entonces podemos decir que la lógica sobre la cual se fundamenta el paso del lenguaje

⁸⁶ MARTINO, A. A. (1985). *La informática jurídica hoy*. Buenos Aires: Ediciones Depalma. p. 556.

⁸⁷ VARGAS OSORNO. T. G. (1990). *Op.cit.*, p. 73.

común (multívoco e impreciso) al lenguaje algorítmico (unívoco y preciso) es la moderna, formal (no dialéctica) y de dos valores⁸⁸.

La parte más elemental de la lógica formal estudia la manera de transformar la equivocidad del lenguaje cotidiano, en un lenguaje unívoco y preciso. Su punto de partida es la proposición, considerada como un todo unitario, y por ello recibe la denominación de "Cálculo proposicional⁸⁹".

Las Proposiciones son aquellas palabras, sintagmas o símbolos que representan algo o que tienen un sentido lógico. La LOGICA FORMAL estudia las proposiciones del lenguaje ordinario desde el punto de vista de su verdad o falsedad, por lo tanto no son objeto de cálculo proposicional aquellas frases de las cuales no se puede predicar la verdad o falsedad. Las proposiciones verdaderas se representan por el símbolo 1 y las falsas por el 0. Valorar una proposición es averiguar si es verdadera o falsa.

Para el cálculo proposicional cualquier frase no constituye una proposición en sentido estricto; por ello lo primero que establece la lógica formal son las reglas mediante las cuales se elaboran las proposiciones fundamentales. De la combinación de estas proposiciones entre sí se ocupa la Sintaxis.

La Sintaxis de un lenguaje simbólico está integrada por reglas de formación y reglas de transformación. Estas reglas son exteriores al sistema que se quiere construir. Las reglas de formación tienen como objeto establecer taxativamente los signos por

⁸⁸ Ibídem. p. 74

⁸⁹ SERRANO PIRELA, A. (1975). *Computadoras y derecho: una introducción a la informática jurídica*. Maracaibo: Universidad de Zulia. Facultad de Derecho, Centro de Estudios de Filosofía del Derecho. p. 28.

utilizar y las posibles combinaciones de los mismos. Los signos indican el objeto al cual se refiere la proposición y se trata de letras minúsculas del alfabeto. Existen también otros signos por medio de los cuales se indican las propiedades del objeto aludido. Su estudio constituye un sector específico de la lógica formal, el cual se conoce como Cálculo de Predicados⁹⁰.

Además de estos signos, existen los signos lógicos, los cuales mediante las reglas de transformación, permiten la construcción de proposiciones complejas que se dividen en:

- Signos Conectivos que son los siguientes:

\wedge = Y: Conjunción $P \wedge Q$ Se lee P y Q

\vee = O: Disyunción $P \vee Q$ Se lee P o Q

\sim = NO: Negación $P \sim Q$ Se lee P pero no Q

\rightarrow = SI: Entonces $P \rightarrow Q$ Se lee P entonces Q

\leftrightarrow = Si Y SOLO SI: Bi implicación $P \leftrightarrow Q$ Se lee P es...si y solo si Q

es... -Signos operadores, denominados cuantificadores y son:

$(\forall x)$ = Para todas las X vale que: Denominado Cuantificador Universal

$(\exists X)$ = Existe al menos una X que: Denominado Cuantificador Existencial

Para determinar el valor de verdad de los enunciados propuestos se utilizan las TABLAS DE VERDAD. Para los conectivos mencionados son:

⁹⁰ Ibídem. p. 29.

CONJUNCION

Sea $P \wedge Q$

P	Q	$P \wedge Q$
1	1	1
0	1	0
1	0	0
0	0	0

Es verdadero solamente en caso de que ambas proposiciones sean verdaderas.

DISYUNCION

Sea $P \vee Q$

P	Q	$P \vee Q$
1	1	1
1	0	1
0	1	1
0	0	0

Es verdadera si al menos una de las proposiciones es verdadera. Únicamente será falsa cuando ambas proposiciones lo sean

NEGACION

$$P \sim Q$$

P	Q
1	0
0	1

Si la primera proposición es verdadera, la segunda es falsa y viceversa

IMPLICACION

$$P \leftrightarrow Q$$

P	Q	$P \leftrightarrow Q$
1	1	1
1	0	0
0	1	0
0	0	1

Es falsa únicamente en el caso de que la primera sea verdadera y la segunda falsa.

1.7.2. Operadores lógicos

Los Operadores son aquellos conectores, por conducto de los cuales se establece una lógica de consulta, relacionando los términos queridos, y creando un simulacro de lo

que será una frase en lenguaje humano, donde los operadores hacen las veces de conjunciones o verbos.

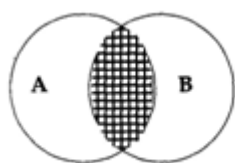
Los Operadores lógicos son elementos informáticos que permiten recuperar la información con inclusión de determinados conceptos y exclusión de otros.

Los Operadores lógicos, son denominados así, por basarse en el álgebra binaria de Boole y permiten que la consulta pueda tener varias palabras, y al sistema realice la búsqueda con base en criterios de su presencia simultánea, alternativa o excluyente.

Los Operadores lógicos de carácter básico son:

- Operadores de Inclusión [Y]: Es conectivo y corresponde a lo que en la lógica de conjuntos se denomina intersección.

$A \cap B$:



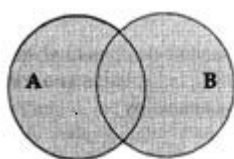
Significa que los términos A y B deben encontrarse en la misma unidad de información, en el mismo documento, para resolver la consulta. Ejemplo:

<Responsabilidad personas plurales> y <Daño>. El empleo de estos dos descriptores unidos con el operador <Y> nos permitirá recuperar todos los documentos en que

aparezcan los dos conceptos jurídicos, cualquiera que sea su orden y se encuentren solos o con otros operadores.

- Operador Alternativo [O]: Corresponde a lo que en la lógica de conjuntos se denomina conjunción.

$A \cup B$:



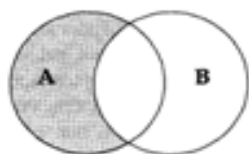
Significa que resuelven la consulta todas las unidades de información en donde se encuentren los términos A o B. Tiene por resultado la suma de ambos.

Ejemplo:

<Compraventa> o <Compraventa inmuebles>

- Operador Exclusivo [NO]: Corresponde a lo que en la lógica de conjuntos se denomina diferencia. Es el inverso del operador <Y>. Exige que esté presente el primer término, pero que no aparezca el segundo en el documento.

$A \cap B$



Significa que resuelven la consulta las unidades de información en las que se encuentra el término A pero no el B, excluyéndose aquellas unidades donde se encuentran en forma simultánea.

Permite recuperar aquellos documentos que contienen los descriptores deseados, excepto los que, a su vez, ostenten los descriptores no deseados.

Ejemplo:

<Responsabilidad en actividades peligrosas> Salvo <Presunción de culpabilidad>

Esto nos permitirá encontrar todos los documentos que se ocupen de la Responsabilidad en actividades peligrosas, excepto los que traten de la presunción de culpabilidad.

Los Operadores lógicos no siempre actúan solos, sino conectados entre sí, lo que permite seleccionar cuestiones cada vez más complejas y al mismo tiempo una selección muy fina

1.7.3. Operadores sintácticos

Cuando se trata de formar una ecuación lógica, no basta que los descriptores aparezcan, sino que el orden entre los mismos es esencial.

Los Operadores sintácticos más utilizados son:

Operador de Vecindad: “Exige que las dos palabras o descriptores se encuentren uno junto al otro, y en el mismo orden en que se solicita o en orden inverso, sino se utiliza el operador de orden”⁹¹.

Ejemplo:

Arrendamientos + Urbanos

Arrendatario + Derechos

Operador de Distancia:

“Establece la distancia máxima a que debe encontrarse una palabra, cuya unión con otra se pretenda. Se expresa con el signo +, y a continuación el número máximo de distancia. Así (+1): significa que ambas palabras deben estar juntas, (+2) que sólo puede haber una palabra de por medio, (+3) que sólo puede haber dos palabras y así sucesivamente”⁹².

Operador de Distancia en Frase o Párrafo: “Exige que las palabras solicitadas se encuentren dentro de un mismo párrafo o de una misma frase, separándose unas a otras por punto seguido, punto aparte o por cualquier signo ortográfico. Se utiliza las abreviaturas WITH para la frase y SAME para el párrafo”⁹³.

Operador de Orden: “Exige que las palabras buscadas estén juntas y en el mismo orden en que se piden”⁹⁴.

⁹¹ VARGAS OSORNO. T.G. op. cit. p. 81.

⁹² *Ibídem*.

⁹³ *Ibídem*. p. 82.

⁹⁴ *Ibídem*.

1.7.4. Operadores de extensión

Los operadores de extensión son aquellos que permiten buscar palabras que puedan tener relación con la deseada, sin tener que expresar todas ellas. Entre ellos tenemos:

Truncadura a la Derecha: Se establece cuál es la raíz de la palabra y se buscan todas las relacionadas con ella.

Ejemplo:

Compra, nos dará: Compra – Dor

Compra – Venta

Truncadura a la izquierda: Se obtendrán todas las palabras que terminen como la pedida.

Ejemplo:

Acción, nos dará Contra – Acción

Re – Acción

Doble truncadura a la derecha - izquierda: Con un ejemplo explicaremos:

- Licit - nos dará: I - Licit – 0

Licit -0

I - Licit – UD

1.8. Elaboración de bancos de datos jurídicos

Entendemos por banco de datos jurídicos, el conjunto de documentos de naturaleza jurídica, bien sea de carácter general o especial, que compila, analiza y permite obtener la información pertinente al tema objeto de búsqueda, estos contienen toda la documentación relevante sobre aspectos netamente jurídicos a saber: Leyes (entendidas en sentido amplio), jurisprudencia y en algunas ocasiones documentos de carácter doctrinal.

Para Emilio Suñé Llinás la base de datos o banco de datos “es el conjunto total de la información que se va a tener disponible”⁹⁵. Así mismo dentro de los bancos de datos encontramos registros, que son cada una de las partes en que la información se divide y los registros a su vez se fragmentan en campos, que son unidades elementales de información⁹⁶. En los bancos de datos los campos se agregan en registros, los registros en archivos, y los archivos en base de datos⁹⁷.

Dentro de la profesión jurídica, encontramos diversidad de especialidades, cada una de ellas con un régimen jurídico propio, que en virtud de los cambios genera la necesidad de actualizaciones constantes. Como una respuesta a la inflación jurídica surge la necesidad de desarrollar bancos de datos jurídicos, que otorgan la posibilidad de recuperar la información, estableciendo sus alcances y límites, de una forma ordenada y ágil.

⁹⁵ SUÑÉ LLINÁS, E. (1994). *Informática práctica para juristas y profesionales del mundo de las letras*. Madrid: Servicio de Publicaciones, Facultad de Derecho, Universidad Complutense y Beramar. p. 184.

⁹⁶ *Ibídem*.

⁹⁷ *Ibídem*.

En la elaboración de un banco de datos jurídico, lo primero que debemos hacer es la de delimitar el tema objeto de estudio, bien sea derecho de familia, civil, tributario, medio ambiente, etc., se deberá entonces catalogar la información y limitarla al conjunto de fuentes jurídicas que regulen exclusivamente esa materia, teniendo siempre en cuenta que en virtud de los criterios de interpretación, se puede recurrir a normas de otras áreas del derecho.

Se deben definir claramente los contornos del contenido del banco y los documentos que van a ser incorporados, es decir, hay que definir el cubrimiento del respectivo banco de datos. Lo ideal es que un banco de datos contenga toda la legislación (leyes, decretos, resoluciones, circulares), y la jurisprudencia (por lo menos la de las altas cortes).

Ahora bien, la existencia de los bancos de datos aplicados a la ciencia del derecho, además de permitir la creación de los mismos en áreas concretas del conocimiento jurídico, también pueden admitir la creación de un banco de datos: global o sectorial. Los bancos de datos *globales*, abarcan la totalidad del ordenamiento jurídico de determinado país, dentro de estos, no existe el problema de definir su contenido, pues éste comprende todas las áreas del derecho. Con relación a ellos sólo se presenta la necesidad de definir la clase de documentos que incluirá, pues podrá referirse sólo a las leyes y decretos con fuerza de ley, o agregar también los reglamentarios; y de éstos no sólo los de alcance nacional, sino también los de carácter departamental y municipal. Lo mismo ocurriría con la jurisprudencia si se incluyera todas las de las altas cortes o si también la de los tribunales superiores, e inclusive la de los juzgados.

En cuanto a la definición de contenido, si se trata de un banco de datos jurídico global, la definición del mismo no presentará mayor inconveniente, esto se encuentra dado por la totalidad del ordenamiento jurídico del país específico donde se pretenda crear la base de datos. Por otra parte, si lo que se busca es crear una base de datos sectorial habrá que definir de forma precisa los documentos jurídicos que deben formar parte del mismo teniendo en cuenta su contenido.

Son *sectoriales* Cuando cubren la legislación, la jurisprudencia y doctrina de un sector definido del derecho (derecho privado, derecho del medio ambiente, contratación estatal etc.), o una de las tres áreas en algún sector. Cuando el banco es sectorial, lo primero que se hace es definir los contornos del mismo, a fin de determinar con claridad el área del derecho a que se referirá, estableciendo límites precisos.

Una vez determinado los contornos del banco de datos que se pretende crear, esto es: global o sectorial, se debe realizar en primera instancia la recopilación de la normativa y en general de las fuentes jurídicas que regulen la materia en particular, resulta necesario que esta se realice de manera minuciosa de forma tal que se incluya la totalidad de los documentos jurídicos, razón por la cual esta primera fase es relevante para efectos de tener una compilación integral y en especial completa.

El segundo paso a realizar después de delimitar el banco de datos es la recopilación de la totalidad de los documentos que contengan las fuentes normativas correspondientes. Con base en ella se hace un análisis sistemático de las distintas

instituciones jurídicas que integran el área del derecho que se van a incorporar al banco, se establecen las relaciones jerárquicas de cada una de ellas determinando el género próximo al cual pertenecen y las especies en que se divide y se define el alcance de los distintos conceptos: En una palabra se adquiere claridad sobre el contenido teórico que se va a sistematizar: Esta labor debe estar dirigida por un abogado especializado en la materia, quien debe continuar la asesoría a los analistas de todo el proceso de elaboración de tesauros y recopilación de la información.

Hecha la recopilación, surge la necesidad de hacer una primera denominación de los temas más generales que se pretenden estudiar, esta primera aproximación general a los temas se le denomina *macrotesauros*, estos varían dependiendo del tipo de tesaurus que se pretende crear, así pues los niveles de la desagregación de la información en un macrotesaurus se encuentra supeditada a la profundidad con la cual se pretenda abordar el banco de datos.

En la *elaboración de macrotesauros*, se tiene en cuenta en primer lugar los estatutos básicos correspondientes, los cuales suelen estar divididos en libros, títulos, capítulos, secciones y apartes. ¿Hasta cuál de estas divisiones debe llegarse para determinar los términos del macrotesaurus? Esto dependerá de los fines del banco que se está elaborando. Por ejemplo si se trata de información legal con fines legislativos, la especificidad tiene que llegar a unidades temáticas concretas, como sería la de cada una de los contratos que trae el Código Civil.

Por lo anterior, los macrotesaurus se pueden crear únicamente con fines de manejo de información bibliográfica, con temas genéricos, verbigracia en derecho tributario,

impuesto sobre la renta, impuesto al patrimonio, cuatro por mil, IVA. Esta primera aproximación se deberá realizar, teniendo en cuenta los encabezados de las leyes.

Por otra parte, “los capítulos tienen nombres que no corresponden a unidades temáticas, como “disposiciones generales”, por lo que es necesario leer su contenido con el fin de observar si contienen temas que merezcan introducirse en el macrotesauros. Igual ocurre con muchas leyes y decretos en donde no aparecen divisiones; en estos se pueden detectar los temas genéricos mediante una lectura general del encabezado y del articulado”⁹⁸.

Como anotamos con anterioridad, el primer paso es fijar el contenido del documento, para lo cual se hace necesario leer el documento y establecer los temas generales contenidos en el mismo, que las más de las veces son los títulos de cada documento legal, no obstante habrán casos excepcionales tales como las disposiciones generales en los cuales habrá que entrar a determinar la existencia de un tema dentro de su contenido.

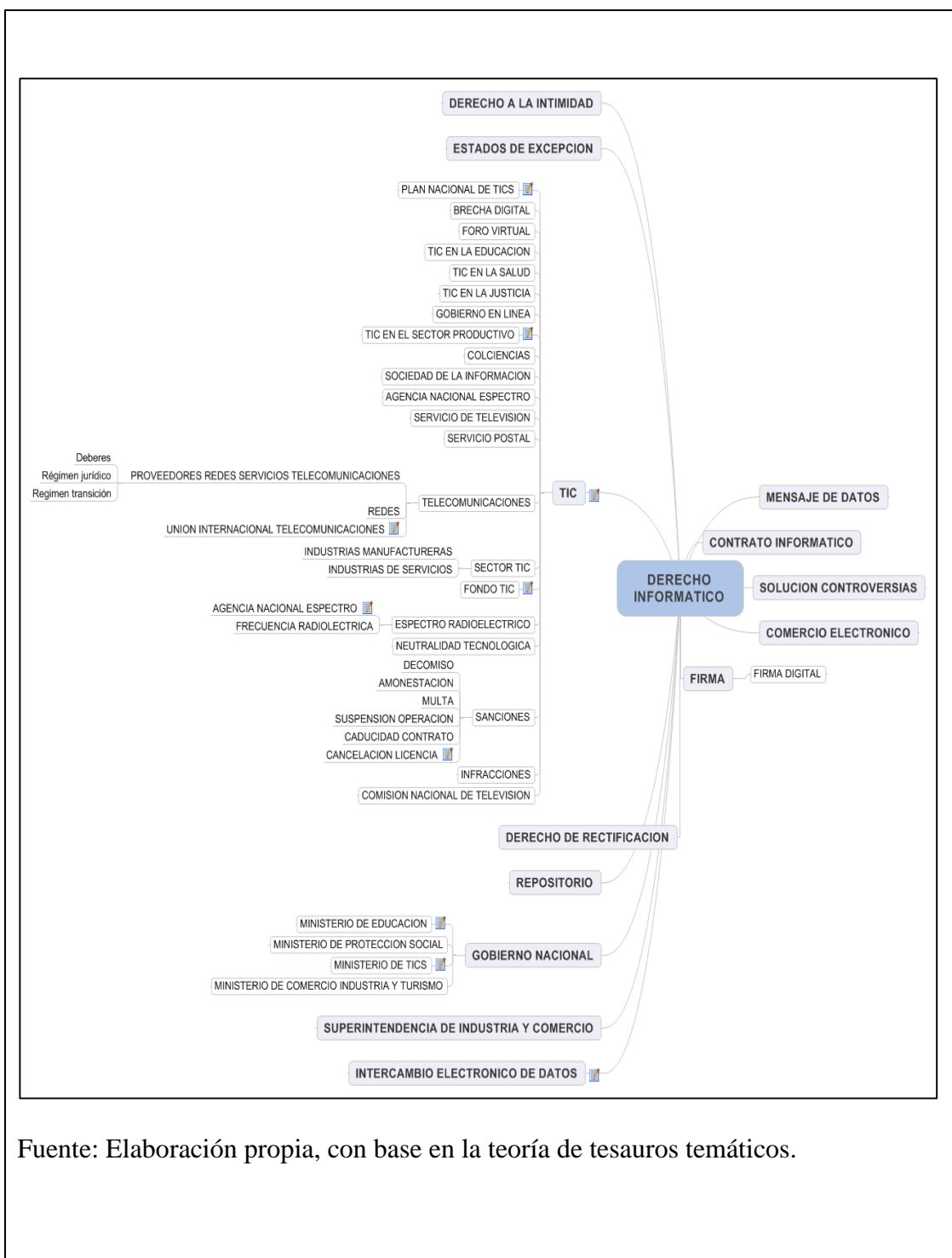
Este primer análisis, no pretende en ningún momento fijar la totalidad de los descriptores que va a contener el tesauo, lo que busca es establecer las bases generales “de la temática a partir de los grandes conceptos que la integran, y sus divisiones hasta el nivel que se haya determinado en forma anticipada, según lo explicamos atrás. Es decir, que la búsqueda de la temática se hace mediante la elaboración simultánea de la correspondiente estructura arborescente.”⁹⁹.

⁹⁸ GIRALDO ÁNGEL, J. (1990). Op.cit., p. 28.

⁹⁹ Ibídem. p. 29.

Para ilustrar mejor el tema, incorporaremos una muestra de un macrotesauro, desarrollado para la creación de un banco de datos jurídicos en derecho informático:

Figura 1. Macrotesauro en derecho infomático.



Fuente: Elaboración propia, con base en la teoría de tesauros temáticos.

Desarrollado el macrotesauro, se empieza a elaborar el tesauro especializado, que se inicia con el análisis minucioso de la lectura de toda la legislación existente de la materia y la asignación de los descriptores, que luego se ubican dentro de la estructura arbórea del macrotesauros en el nivel de desagregación de la información que corresponda, hasta llegar a niveles especializados de información.

Desarrollado el tesauro especializado, el siguiente paso es el análisis del documento jurídico que se pretende incorporar a la base de datos. Ahora nos detendremos en cada uno de los documentos, con el fin de saber cómo debemos analizar esta información, con fines de ingresarla al banco de datos, tarea que constituye la base fundamental para su elaboración, y que podrá ser el soporte para el jurista.

El contenido de un banco de datos jurídicos, centra su atención principal y fundamentalmente en tres documentos a saber: El documento legal, el documento jurisprudencial y el documento doctrinal. Cuando mencionamos el documento legal, no hacemos referencia a su sentido formal, aquí la ley contempla leyes, decretos, resoluciones, circulares externas, etc.

Es importante mencionar que el concepto de bancos de datos jurídicos ha sido definido, por Emilio Suñé Llinás como aquellas que contienen datos propiamente jurídicos. Las normas, son datos jurídicos en todos los sistemas junto a la jurisprudencia, al ser las únicas que contienen datos jurídicos en sentido propio.

El profesor Suñé no está de acuerdo, opinión a la cual me sumo, contrario a la mayoría de los autores, cuya perspectiva es mucho más general, en la que le atribuyen el carácter de bases de datos jurídicas a las que recogen el contenido de las publicaciones de los parlamentos, estas sin duda, tiene un diáfano interés para la interpretación de las normas, pero lo mismo se la puede considerar jurídica que de ciencia política. O aquellas otras, que contienen bibliografía jurídica, es importantísimo para nuestra disciplina; pero no se hace de forma distinta que otra base de datos bibliográfica sobre literatura o astronomía; no en vano existen normas idénticas para la informatización de todo tipo de bibliografía, de las cuales la más usada es el formato MARC, diseñado para la biblioteca del Congreso de los Estados Unidos y cuya adaptación española se denomina IBERMAC. Si se aceptara tanta generalidad se estaría permitiendo difuminar totalmente los límites de la informática jurídica.

Así pues encontramos que dentro del documento legal, “la unidad informática es el artículo, cualquiera que sea la extensión, el número de incisos que lo conformen, e independientemente de que tenga o no párrafos”¹⁰⁰.

Con el fin de fijar los descriptores de todos y cada uno de los artículos del documento, doctrinalmente se ha fijado una metodología que permite hacer mucho más objetivo el procedimiento dentro del documento legal, a saber:

1. “Se determinan en cada artículo los términos jurídicos que tengan contenido conceptual autónomo.

¹⁰⁰ Ibídem. p. 53.

2. Se agregan los implícitos. En algunas ocasiones en el artículo no aparecen algunos de los términos a que se refiere el contenido del mismo, por lo que será necesario agregarlos para que la información pueda ser recuperada por conducto de ellos, es lo que se denomina el *análisis de los implícitos*¹⁰¹. Verbigracia:

Ley 100 de 1993, art. 72.

“Cuando el afiliado se invalide sin cumplir con los requisitos para acceder a una pensión de invalidez, se le entregará la totalidad del saldo abonado en su cuenta individual de ahorro pensional, incluidos los rendimientos financieros y adicionado con el valor del bono pensional si a ello hubiere lugar.

No obstante, el afiliado podrá mantener un saldo en la cuenta individual de ahorro pensional y cotizar para constituir el capital necesario para acceder a una pensión de vejez”¹⁰².

Del texto del artículo transcrito no se desprende directamente a qué acciones se refiere, por lo que será necesario consultar otras normas del decreto que permitan determinarlas, para incluirlas como descriptores.

3. “Se descartan los que tengan contenido ajeno a la temática del banco, teniendo en cuenta que se trata de un banco especializado en un área delimitada del derecho”¹⁰³.

¹⁰¹ Ibídem.

¹⁰² CONGRESO DE LA REPÚBLICA DE COLOMBIA. (1993, diciembre 23). *Ley 100, por la cual se crea el sistema de seguridad social integral y se dictan otras disposiciones*. Diario Oficial, No.41.148.

¹⁰³ GIRALDO ÁNGEL, J. (1990). Op.cit., p. 53.

A diferencia de los bancos de datos especializados, los bancos de datos globales, que no representan mayor complejidad, dado que únicamente se ubicará el libro, título y capítulo donde reposa la información objeto de búsqueda.

4. “Se seleccionan los términos para los cuales haya información relevante en el texto del artículo, y se descartan los demás¹⁰⁴”.

5. “Determinadas las palabras que deben asignarse como descriptores, se buscan en el tesoro los términos que correspondan y se anotan en el espacio correspondiente de la *hoja de entrada*”¹⁰⁵.

Una vez estudiado el documento legal, nos corresponde explicar el *documento jurisprudencial*. Si en el documento legal la unidad informática era la norma, se ha afirmado que en este caso la unidad informática es la sentencia¹⁰⁶. Esto se encuentra fundamentado principalmente en que la única forma de poder incluir un descriptor sobre una sentencia es analizando el contenido de la misma, en particular creando y resolviendo el o los problemas jurídicos que se plantean dentro de la misma.

Existen dos sistemas de análisis jurisprudencial, cuya incidencia depende del sistema que rija al país, esto es del *common law* o *civil law*. Así, si nos encontramos en el *comon law*, resultara necesario e imprescindible la incorporación del análisis fáctico que describa de manera detallada las variables, con el fin de tenerlas presentes para

¹⁰⁴ Ibídem. p. 54.

¹⁰⁵ Ibídem.

¹⁰⁶ Ibídem. p. 57.

una aplicación futura en virtud del precedente. A diferencia del *civil law*, donde se deberá analizar las consideraciones jurisprudenciales con el fin de fijar los alcances e interpretaciones normativas, de forma general y abstracta.

En el análisis jurisprudencial del *civil law*, se podrá prescindir de incluir la totalidad del texto jurisprudencial, aunque como veremos más adelante, existen bancos de datos jurídicos, como es el caso de los bancos de datos de la Universidad Externado de Colombia, que incluyen la totalidad de la providencia judicial, planteamiento del problema jurídico, tesis y su análisis pertinente con los correspondientes descriptores, con el fin de suministrarle al usuario la información más completa. No obstante lo anterior, hay quienes sostienen que no tiene sentido incorporar el texto de la providencia¹⁰⁷.

En nuestro sentir, resulta conveniente realizar la incorporación total del documento jurisprudencial, entre otras, por la razón anteriormente expuesta, esto es, suministrar la información más completa posible y fundamentalmente, debido al subjetivismo en el cual podemos caer al realizar un análisis resumido del documento, en especial, porque se corre el riesgo de no incluir aspectos que para el usuario si resulten necesarios, de manera que es conveniente tener el documento original dentro de la base de datos.

¹⁰⁷ Ibídem. p. 59. "Coincidimos con la idea de que en los países de legislación codificada no tiene sentido incorporar el texto completo de la providencia. Inclusive vamos más allá al considerar que muchas de las partes los dos bancos atrás mencionados deben también eliminarse: En primer lugar, los hechos concretos que permiten la formulación de la doctrina jurisprudencial de carácter general y abstracto. En segundo lugar, tampoco se deben incorporar los resúmenes que se hacen en la providencia analizada de las decisiones de instancia y las argumentaciones de las partes, pues si ellas tienen valor jurídico que amerite su incorporación al banco, ello se debe hacer con base en el documento original."

El sistema más utilizado en la elaboración de bancos jurídicos que contienen documentos jurisprudenciales, es el de realizar la inclusión del texto jurisprudencial, independientemente de si se trata de un sistema *common law* o un sistema *civil law*. Aclarando en este punto, que como dijimos anteriormente, el análisis varía según el sistema que rija el país, aquí de lo que se trata es de la incorporación del documento.

Una vez incluido el documento jurisprudencial dentro de la base de datos, corresponde formular los problemas jurídicos que suscite la sentencia, pueden existir tantas tesis como problemas jurídicos, ya que la tesis es la respuesta al problema y son consideradas la parte más relevante a tener en cuenta sea en uno u otro sistema, dado que son de obligatorio cumplimiento. Los problemas jurídicos deberán formularse de manera general y abstracta en el sistema *civil law*.

En este mismo sentido, es importante recordar que dentro del sistema *civil law*, las tesis que dan respuesta a los problemas jurídicos planteados, igualmente deben ser de carácter general y abstracto, por lo cual no podrán contener particularidades de los aspectos fácticos, toda vez que no tienen aplicación. Diferente será el análisis jurisprudencial en el sistema *common law*, donde deberá incluirse el fundamento fáctico con el fin de tener en cuenta el precedente.

Ahora bien, dentro del análisis del documento jurisprudencial, deberán estar presentes los criterios de interpretación, es decir la aplicación de la hermenéutica, bien sea teniendo en cuenta el método exegético, sistemático y sociológico, etc. La aplicación de uno u otro método de interpretación pueden generar la creación de soluciones distintas al problema jurídico planteado, de ahí que lo que se deberá

realizar es la incorporación de “los juicios hermenéuticos que sirvieron para determinar el contenido y alcance de las normas”¹⁰⁸.

Algunos autores, mencionan la existencia de cuatro casos, en los cuales hay lugar a incluir la providencia en el banco de datos, estos son:

“Tesis sustentadas en la prueba de los hechos. Las tesis que se sustentan en el análisis de la prueba no pueden aplicarse a casos diferentes, porque los fenómenos fácticos son irrepetibles¹⁰⁹”.

“Tesis sustentadas con base en la connotación jurídica de los hechos. Cuando las tesis se orientan a afirmar que con base en determinados hechos muy particulares del caso en estudio se infiere un específico fenómeno jurídico¹¹⁰”

“Problemas cuya solución se funda en el fenómeno jurídico. Por ser el negocio jurídico una fuente que solo obliga a las partes que intervienen, los problemas jurídicos que se resuelvan con base en él no dan lugar tampoco a documento jurisprudencial. Sin embargo, cuando se trata de un contrato o una convención colectiva de trabajo que involucra a un número significativo de personas, las decisiones que se tomen con base en ellos pueden servir para resolver problemas similares de personas que también sean parte de ellos por lo que tales decisiones tienen interés para bancos especializados en donde ellos pueden tener eficacia jurídica.¹¹¹”

¹⁰⁸ Ibídem. p. 88.

¹⁰⁹ Ibidem. p. 105.

¹¹⁰ Ibídem.

¹¹¹ Ibídem. p. 108.

“Problemas en los que sólo se enumera la tesis sin fundamentarla jurídicamente.

La naturaleza del trabajo de los despachos de instancia (juzgados y tribunales) suele orientarse fundamentalmente a la aplicación del derecho, por oposición al trabajo de la Corte Suprema de Justicia, en donde prima el análisis de los conceptos jurídicos sobre la solución misma del caso concreto.”

Por esta razón es muy frecuente encontrar en las sentencias de instancia, tesis jurídicas con relación a las cuales no se hace el análisis ni de su vigencia ni de su alcance, sino que se toman como “a priori” no discutibles, respecto de los cuales solo se trata de determinar si en el proceso en estudio aparecen”¹¹².

Para concluir la explicación acerca del análisis jurisprudencial, a la tesis que da respuesta al problema jurídico deberá asignársele uno o varios descriptores, con el fin de organizar la información y facilitar su búsqueda. Es relevante mencionar que la inclusión de descriptores permite la recuperación de información, sin embargo, no es la única forma, recordemos que los datos identificadores del documento son criterio de búsqueda de información.

En cuanto a la asignación de descriptores, estos siguen los mismos pasos establecidos para el documento legal. Existen bancos de datos que contienen documentos jurisprudenciales, dentro de los cuales no se incluyen los descriptores, estos no son otros, que los bancos de datos globales, ya explicados anteriormente. Un

¹¹² Ibídem. p. 112.

ejemplo palmario; es la base de datos de la Corte Suprema de Justicia de Colombia, dentro del cual carece de sentido la inclusión de los descriptores, debido a la naturaleza jurídica del mismo. No obstante, lo propio no ocurre con los bancos de datos especializados, en los cuales si se deberán incluir.

Otro aspecto que se deberá tener en cuenta, con el fin de ofrecer información completa sobre un tema específico, consiste en realizar la correspondiente referencia a sentencias en igual sentido y sentido contrario. Esta asignación se realiza con base en las citas que se incluyen dentro del documento jurisprudencial sobre otras sentencias emitidas por la misma corporación o por otra diferente. Este procedimiento, proporciona una recuperación de información integral que podrá estudiar otras ópticas y profundizar sobre los temas contenidos en la jurisprudencia.

Puntualmente, existen otros aspectos que se deberán tener presentes al momento de realizar un análisis jurisprudencial. Esto es, la inclusión de las normas a que se hagan referencia dentro de la sentencia. Así pues, cuando se pretenda realizar la recuperación de información de una norma, esta hará referencia a la existencia de jurisprudencia en la materia e igualmente se podrá consultar la norma para comprender en mayor medida el contenido de la jurisprudencia.

Sobra mencionar que los datos identificadores de la sentencia deben encontrarse presentes en la incorporación del documento en el banco de datos, siendo base fundamental para la obtención de información.

Dentro del ámbito jurídico, al igual que encontramos el documento legal y jurisprudencial, dentro de algunas bases de datos de incluyen los *documentos doctrinales*, estos, son escritos jurídicos que realizan un aporte significativo al área del derecho objeto de análisis, bien sean obras o ensayos que estudian un tema jurídico en particular, esta clase de documentos, cuenta al igual que las otras categorías, con un procedimiento metodológico para su análisis.

La unidad informática del documento doctrinal, es cada uno de los artículos, capítulos, en los cuales se encuentra dividido, de tal forma que una vez establecida cada temática se procede a la asignación de los descriptores correspondientes, teniendo en cuenta el tema central que trata cada una de las unidades.

Aquí, es importante mencionar que el análisis del documento doctrinal se asemeja al del documento jurisprudencial, dentro del cual pueden surgir varios problemas jurídicos con sus respectivas tesis, resulta pertinente incorporar la totalidad del texto objeto de análisis dentro del banco de datos, garantizando que cada unidad contenga un tema jurídico en particular, de tal forma que se evite incluir escritos innecesarios que no correspondan al problema jurídico planteado.

En cuanto a la asignación de descriptores, estos se deberán asignar a partir de la tesis formulada seleccionando los términos adecuados que reposan en el tesauro que ha sido elaborado previamente. Es importante mencionar aquí, que los términos son de contenido metajurídico, de manera que solo se podrán asignar descriptores

contenidos en el tesoro, pues de lo contrario habría que romper la estructura de este para incorporar términos de ciencias no jurídicas¹¹³.

1.9. Panorama internacional

La existencia de los bancos jurídicos en el mundo, cuenta con grandes dimensiones técnicas. Por esta razón, presentaremos los más importantes, a manera enunciativa, reconociendo sus aspectos más importantes.

1.9.1. Unión Europea

La Unión Europea¹¹⁴, cuenta con bases de datos jurídicas sobre legislación: directivas, reglamentos y decisiones en curso de adopción o que han entrado en vigor en el territorio europeo.

De igual forma, cada institución europea cuenta con un portal de internet de acceso gratuito, es así, como el Parlamento Europeo en su sitio web cuenta con un hipervínculo que ofrece una base de datos jurídica sobre sus archivos¹¹⁵, en esta página es posible consultar los procedimientos legislativos ordinarios, de aprobación, los procedimientos no legislativos, de cooperación, presupuestarias, de aviso de conformidad, los procedimientos institucionales sobre la nominación o la moción de censura, entre otros.

¹¹³ Ibídem. p. 147.

¹¹⁴ UNION EUROPÉENNE. (2014, enero 18). Site web officiel de l'Union européenne. Législation de l'UE: http://europa.eu/eu-law/index_fr.htm

¹¹⁵ PARLEMENT EUROPÉEN. (2014, enero 18). Archives: <http://www.europarl.europa.eu/parlArchives/comArch/comDocSearch.do;jsessionid=B91E98AD4078138EE6655FB2382A475E.node1?language=fr>

De igual forma a nivel presidencial, el Consejo Europeo cuenta con su sitio de internet dentro del cual es posible tener acceso a reuniones ordinarias y extraordinarias del mismo¹¹⁶.

A nivel ministerial se encuentra el Consejo de la Unión Europea, que contiene un registro de todos los documentos desde el 1 de enero de 1999 por un sistema de archivo automático¹¹⁷.

Así, todo documento no sensible sometido al Consejo o a una de sus instancias preparatorias, que pueda servir de base a las deliberaciones, como puede tener una influencia sobre el proceso de la decisión o reflejar los progresos logrados sobre un asunto, es automáticamente mencionado en el registro.

De otra parte dentro de las instituciones europeas, se encuentra el sitio web de la Comisión Europea¹¹⁸, en este sitio es posible tener acceso a la legislación adoptada y la que se encuentra en curso de elaboración. En lo relacionado con la jurisprudencia europea, es la Corte de Justicia de la Unión Europea la que ofrece una compilación de las decisiones del Tribunal de la Función Pública, los tribunales y de la Corte de Justicia de la Unión Europea en su sitio de internet¹¹⁹.

Finalmente existe un sitio de la Unión Europea denominado EUR-Lex, en este portal web se reúnen todas las bases jurídicas del derecho europeo, a saber: los

¹¹⁶ CONSEJO DE LA UNIÓN EUROPEA. (2014, enero 18). <http://www.consilium.europa.eu/es/council-eu/>

¹¹⁷ *Ibíd.*

¹¹⁸ COMMISSION EUROPÉENNE. (2014, enero 18). http://ec.europa.eu/legislation/index_fr.htm

¹¹⁹ TRIBUNAL DE JUSTICIA DE LA UNIÓN EUROPEA. (2014, enero 18). http://curia.europa.eu/jcms/jcms/j_6/

reglamentos, directivas, decisiones, los casos de la Corte de Justicia de la Unión Europea.

En la Unión Europea la base de datos CELEX (Communitatis Europaeae Lex) Tendencias en la investigación sobre recuperación de información jurídica (Rev. Esp. Doc. Cient., 26, 2, 2003 193 Lex) nace en 1966 y está operativa desde 1970, se empleó el software de recuperación MISTRAL e incorpora además de legislación comunitaria, jurisprudencia del Tribunal de Primera Instancia y del Tribunal de Justicia de las Comunidades Europeas, Trabajos preparatorios, Disposiciones Nacionales de Ejecución de las Directivas Comunitarias y Preguntas parlamentarias¹²⁰.

1.9.2. Bélgica

En este país, se existe el banco de datos jurídico CREDOC (Centre de Documentation Automatique du Droit). Creado en 1967 en Bélgica, nacido por la conjunción del trabajo de abogados y notarios y con participación de la Magistratura. Es uno de los sistemas más desarrollados de Europa, constituyendo la primera experiencia con éxito. Es el sistema de recuperación de información pionero en el viejo continente y el primero en enfrentarse a la problemática de la documentación bilingüe, ya que cuenta con la existencia de un tesoro Bilingüe (Francés-flamenco)¹²¹, CREDOC utiliza un sistema referencial, y no el sistema del texto integral.

¹²⁰ ALVITE DÍEZ, M. L. (2003). *Tendencias en la investigación sobre recuperación de información jurídica*. Revista Española de Documentación Científica, 26(2). doi:10.3989/redc.2003.v26.i2.137.p. 192-193.

¹²¹ VARGAS OSORNO T.G. ob. cit. p. 83.

“La escogencia del sistema de clasificación ha sido llevada sobre la confección de un diccionario palabras- claves de 4.000 términos. El servicio mínimo comprende:

- La respuesta a todas las cuestiones que conciernen a los textos legislativos, las referencias de doctrina y jurisprudencia.
- Mantenimiento al día de la jurisprudencia.
- Comparación automática de diferentes legislaciones.¹²²”

1.9.3. Canadá

Centro de Documentación de la Universidad Laval de Montreal, esta universidad ha registrado sobre microfilms indexados las resoluciones y proyectos de ley que se preparan para ser sometidos bien al gobierno federal, bien al gobierno de Quebec.

El Centro de Documentación de la Universidad registra sobre microfilms alrededor de 5.000 páginas por día, con colaboración de la facultad de derecho y de la biblioteca de la universidad. Utilizan el procedimiento Miracode que permite buscar y localizar en pocos minutos cualquier documento solicitado a partir de microfilms.

Uno de los objetivos de este centro de documentación es el intercambio de microfilms indexados con otras universidades, para evitar la repetición de

¹²² Ibídem.

*trabajos y realizar así, una cooperación internacional en materia de información documental.*¹²³

1.9.4. España

*“El inicio de la aplicación de la informática en el contexto jurídico español viene situándose en el último bienio de la década de los sesenta, con la puesta en marcha del Proyecto Ibertraten 1968 y el Plan General de Informática Jurídica entre cuyos objetivos se hallaba el de facilitar búsquedas y recuperaciones jurisprudenciales”*¹²⁴.

- **Antrid**

*Sistema de análisis numérico para el tratamiento y recuperación de la información documental. Lo tiene funcionando el Ministerio de Justicia. Es un sistema de carácter universal que puede utilizarse para analizar toda clase de información documental: No sólo de naturaleza jurídica sino sobre cualquier otra materia y ello tanto para el tratamiento y recuperación de información como también con fines de investigación lingüística, histórica o literaria de textos.*¹²⁵

¹²³ Ibídem. p. 84.

¹²⁴ ALVITE DÍEZ, M. L. (2003). op.cit. p. 26.

¹²⁵ VARGAS OSORNO T.G. ob. cit. p. 84.

- **Ibertrat**

Dependiente del Ministerio Español de Asuntos Exteriores que entró en funcionamiento en 1976¹²⁶

- **Sinade**

Sistema realizado por el Colegio de Abogados de Barcelona¹²⁷.

1.9.5. Estados Unidos

- **Sistema Lexis**

El origen de LEXIS en Estados Unidos se remonta al año 1964, fruto de una iniciativa de la *Ohio Bar Asociación* quien contrata a *Data Corporación* de Dayton (Ohio) para desarrollar un sistema de recuperación. En 1968 fue adquirida por *Mead Corporación*, naciendo *Mead Data Central* que lanzó un prototipo del sistema operativo al año siguiente. El nuevo servicio apareció en 1973 bajo el nombre de LEXIS preocupado, de modo especial, por la documentación judicial y por la creación de un sistema de recuperación fiable y robusta que, incluso es “creado por la Mead Data Central en 1963¹²⁸”.

“Es el banco de Datos Jurídico más grande del mundo, con casi 30 billones de caracteres cargados y con una entrada anual en memoria de alrededor de

¹²⁶ *Ibídem.* p. 85.

¹²⁷ *Ibídem.*

¹²⁸ ALVITE DÍEZ, M. L. (2003). op.cit. p. 191.

35 mil sentencias memorizadas en texto integral. Sistema de carácter privado”¹²⁹.

En la actualidad,

“[...] mantiene muchas semejanzas con la versión inicial. Su principal competidor se desarrolló en el seno de la editorial jurídica más importante de Estados Unidos, West Publishing que sacó a la luz en 1975 su sistema WESTLAW. El software de recuperación inicial, sobre el que se realizaron desarrollos profundos, fue adquirido a la empresa canadiense QL Sistemas. En Canadá, el desarrollo guarda relación con el proceso de Estados Unidos en virtud del uso inicial del software QL Sistema, originariamente denominado QUIC/LAW que nace en 1968 de la mano del profesor Hugh Law ford. El aspecto más singular del sistema residía en que la recuperación se basaba en algoritmos de ordenación que atendían a la frecuencia de aparición de las palabras. En 1974 se prefirió dar prioridad a los conectores booleanas y a los operadores de distancia, reduciendo a simples opciones los algoritmos de ordenación”¹³⁰.

“Lexis es un sistema interactivo de texto completo. Como el sistema es interactivo el abogado puede mantener un diálogo continuo con el ordenador, por medio de un terminal”¹³¹.

¹²⁹ VARGAS OSORNO T.G. ob. cit. p. 86.

¹³⁰ ALVITE DÍEZ, M. L. (2003). op.cit. p. 192.

¹³¹ VARGAS OSORNO T.G. ob. cit. p. 86.

- **Sistema *Iuris***

“Comenzó en 1970 cuando el Departamento de Justicia inició investigaciones y comenzó a desarrollar esfuerzos tendientes a la preparación de un sistema automatizado para recuperar información jurídica. Sistema de carácter oficial”¹³².

“Iuris es un sistema que hace posible la recuperación del texto completo (full-text) de los documentos necesarios para una investigación jurídica, a partir de una fuente central de información. Se trata de un sistema on line, es decir, la información puede ser recuperada del archivo de documentos por medio de una llamada telefónica”¹³³.

- **Sistema *Flite***

“Sistema de carácter oficial. Antiguo Lite (Federal Legal Information Through Electronics), desarrollado a mediados de los años 60 por la fuerza aérea. Este sistema opera con una modalidad Batch, esto es, solamente puede ser satisfecha y atendida una solicitud en cada oportunidad. Sus archivos incluyen el United State Code, las decisiones de la Corte de Apelaciones de los Estados Unidos, los reportes de la Corte Marcial y manual para cortes marciales, tratados y acuerdos internacionales de especial interés para el Departamento de Defensa”¹³⁴.

¹³² *Ibíd.* p. 85.

¹³³ *Ibíd.*

¹³⁴ *Ibíd.* p. 86.

- **Centro de Law Research Service**

*“Comienza en 1964. Utiliza abstract. El sistema parte de una clasificación por materias para cada una de las cuales se ha elaborado un thesaurus que ordena los diversos materiales memorizados numéricamente. El usuario está conectado a una red de télex.”*¹³⁵.

- **Antitrust Division of the Departament of Justice**

Fue el primer servicio de documentación electrónica de textos jurídicos que entró en funcionamiento en los Estados Unidos. Comienza en 1962.

*“Funciona con un sistema de abstract y comprende desde las nociones jurídicas más generales a las más precisas, siempre en materias relacionadas con la jurisprudencia, legislación y doctrina antitrust”*¹³⁶.

1.9.6. Italia

- **Centro de Documentación Automática**

“Los primeros trabajos comenzaron en 1963 bajo la dirección de Angelo Gallizia, quien con la cooperación del ingeniero Enrico Maretti y la abogada Flora Mollame iniciaron el primer Banco de Datos de contenido jurídico destinado al auxilio de la actividad notarial, en Milán.

¹³⁵ Ibídem.

¹³⁶ Ibídem. p. 87.

*El Centro per la Documentazione Automatica comprende sentencias, resoluciones ministeriales y trabajos doctrinales en materia tributaria.”*¹³⁷

- **Centro de Iuscibernética de Turín**

Entre los sistemas más representativos, es obligatorio mencionar el sistema ITAL-IURE FIND surgido de la iniciativa de la Corte Suprema de Casación italiana, el Congreso de los Diputados y el Instituto per la Documentazione Giuridica de Florencia, este último caracterizado hasta el momento actual por el desarrollo de actividades e investigaciones en el terreno de la Documentación, lenguaje jurídico, tecnologías de la información aplicadas al Derecho, etc. Nace en 1968, con la idea de trabajar la jurisprudencia de la Corte de Casación Italiana. Comprende legislación, jurisprudencia y doctrina (tratada por el instituto para la documentación jurídica – IDG- de Florencia) como también documentación sobre contratos colectivos nacionales y datos de los abogados y procuradores. “Funciona en la Universidad de Turín, creado bajo la dirección de MARIO G. LOSANO, desde 1969. El centro se dirigía a impulsar el estudio de la informática en el seno de las facultades de derecho”¹³⁸.

- ***Instituto per la Documentazione del Consiglio Nazionale Delle Ricerche***

“Funciona desde 1968. Posee un Banco de Datos Jurídico constituido por las referencias bibliográficas extraídas de las revistas jurídicas, filosóficas y

¹³⁷ Ibídem.

¹³⁸ Ibídem. p. 88.

de politología. El centro ha compilado un thesaurus de palabras claves seleccionadas a partir de los índices analíticos de documentos doctrinales”¹³⁹.

1.9.7. Francia

En Francia, la iniciativa de la creación de bancos jurídicos partió del Consejo de Estado y el Tribunal de Casación en 1967, “jugando a partir del año 1970 un papel fundamental el CDIJ, posteriormente CNIJ Centre National d’Informatique Juridique. Además de la profusión de sistemas más o menos especializados, se introdujo un sistema de información legal francés basado en el sistema LEXIS, en el año 1983. Como señala Moreno de la Fuente, la fecha importante en el desarrollo de la informática jurídica en Francia, corresponde al año 1985 gracias al denominado «Informe Leclerc», que daría origen al sistema JURIDIAL, eligiendo Questel Plus como software de recuperación”¹⁴⁰.

- **Credij**

(Centre de Recherche et de Development en Informatique Juridique). Creado en 1970. El Centre de Recherche et de Development en Informatique Juridique, unido al Ministerio de Justicia Francés elaboró un sistema de tratamiento automático de documentación jurídica denominado Docilisque comprende una metodología y técnicas documentarias e informáticas adaptadas a la constitución y difusión entre los profesionales del derecho de

¹³⁹ Ibídem.

¹⁴⁰ ALVITE DÍEZ, M. L. (2003). op.cit. p. 192.

*un Banco de Datos que comprende el conjunto de los textos jurídicos franceses. Permite encontrar, con la ayuda del computador, el conjunto de normas, jurisprudencia y doctrina aplicable a una situación determinada.*¹⁴¹

- **El Cridon**

*“(Centre de Recherche D’Informatique et de Documentation Notariales). Creado en 1962, en Lyon. Facilita la documentación deseada y redacta consultas pedidas por los notarios. Sigue el sistema de palabra clave”.*¹⁴²

- **Iretij**

*“(Institute de Recherche et D’études Sur le Traitement de l’Information Juridique). Bajo la dirección del doctor Pierre Catalá. La función de este instituto ha sido primordialmente el estudio de tratamiento de la jurisprudencia francesa. El profesor Catalá, ha sido uno de los primeros que se han ocupado de este tema, trabajando en colaboración la facultad de derecho y la de ciencias de Montpellier. Funciona con un sistema de abstracts”*¹⁴³.

¹⁴¹ VARGAS OSORNO T.G. ob. cit. p. 88.

¹⁴² Ibídem. p. 89.

¹⁴³ Ibídem.

1.9.8. Alemania Federal

- **Juris**

“(Sistema de Información Jurídica). Sistema alemán del Ministerio de Justicia. Es un sistema on line, que se ocupa sobre todo del derecho fiscal y de protección social. Funciona sobre máquinas Siemens, con el software ordinario, Golem-passat, pero con un agregado para el tratamiento lingüístico de los vocablos contenidos en el banco, que se llama Parat. La documentación se refiere al derecho tributario, a la responsabilidad civil y a la previsión social”¹⁴⁴.

- **Datev**

“Sistema de desarrollo privado. Especializado en el campo del derecho fiscal. Contiene documentación en legislación, jurisprudencia, lineamientos administrativos y doctrina”¹⁴⁵.

1.9.9. Brasil

- **Prodase**

*(Centro de Informática y Procesamiento de Datos del Senado Federal).
Creado en 1971, sin embargo su reconocimiento oficial es el 12 de octubre*

¹⁴⁴ Ibídem.

¹⁴⁵ Ibídem. p. 8

de 1972, con el objeto de poder atender al poder legislativo, modernizando y dinamizando la actividad parlamentaria y administrativa del Senado Federal. En 1972, el proyecto del Senado Federal se hace extensivo a la Cámara de diputados. PRODASEM opera en estrecho contacto con el servicio oficial de procesamiento de datos SERPRO. También existen vinculaciones con el poder judicial.

PRODASEM está creando el SIL (Sistema de Información Legislativa) destinada a proveer a los usuarios de los documentos legislativos memorizados.¹⁴⁶

1.9.10. Reino Unido

En el Reino Unido, “los primeros experimentos conducidos por Colin Taper, se remontan al año 1963. Entre 1968 y 1969, se diseñó el SRI legal STATUS en U.K. Atomic Energy Authority en Harwell. Sin embargo, la puesta en marcha de servicios comerciales no llegaría hasta el año 1978, momento en que la editorial Butterworths anunció su acuerdo con Mead Data Central para ofrecer el sistema LEXIS”¹⁴⁷.

¹⁴⁶ *Ibíd.* p. 90

¹⁴⁷ ALVITE DÍEZ, M. L. (2003). *op.cit.* p. 192.

1.9.11. Perú

El profesor Salazar Cano realizó una investigación dirigida a la aplicación del procesamiento electrónico de datos al juicio ordinario con el fin de lograr una mayor celeridad y eficacia en la administración de justicia.

La Universidad Nacional Mayor de San Marcos, a través de la sección de Informática Jurídica y de Derecho Cibernético del Centro de Investigaciones en informática aplicada a las ciencias humanas y sociales, trabaja con estos temas¹⁴⁸.

1.9.12. Venezuela

- **Sistema Juribanc**

El CEFD, adscrito a la facultad de derecho de la universidad de Zulia se encuentra abocado a la organización una base de datos jurídica, designada convencionalmente con la denominación de JURIBANC. Su objeto es lograr La información en los sectores legislativos, jurisprudenciales y de doctrina nacional.

El diseño del proyecto comenzó en el mes de enero de 1973, una vez creado el Centro de Estudios de filosofía del Derecho (CEFD). El proyecto ha sido dividido en dos etapas:

¹⁴⁸ Ibídem.

JURIBANC I: Limitado al almacenamiento de los resúmenes de sentencias, especialmente en materia civil, mercantil y laboral, tanto de algunos tribunales de instancia y superiores como de la Corte Suprema de Justicia. Esta etapa del proyecto tiene el carácter de ensayo-guía, actualmente en curso.

JURIBANC II: Se refiere a la base de datos jurídica en su totalidad. En esta etapa se dio especial tratamiento a la sistematización y almacenaje de la jurisprudencia de la Corte Suprema de Justicia.¹⁴⁹

- **Asociación Latinoamericana de Informática Jurídica y Derecho Cibernético**

“Creada en enero de 1976, en la ciudad de Valencia. Tiene por objeto la realización de actividades de estudio, investigación y aplicación de la Informática Jurídica, el Derecho Cibernético y disciplinas afines en todos los países de Latinoamérica”¹⁵⁰.

1.9.13. México

- **Sistema Unam Iure**

Realizado por el Instituto de Investigaciones Jurídicas de la UNAM, a través de su centro de información de legislación y jurisprudencia. Se creó en 1981,

¹⁴⁹ *Ibídem.* p. 91

¹⁵⁰ *Ibídem.*

siendo director del instituto el doctor Jorge Carpizo, quien en 1982 firma un acuerdo de cooperación entre el Instituto de Investigaciones Jurídicas y el IRETIJ (Institut de Recherche et D'études Pour le Traitement de L'Information Juridique) de Montpellier, encaminada a la creación de un banco de información legislativa en México. La realización práctica del proyecto ha estado a cargo del licenciado Claude Belair Mouchel, jefe del centro, y del ingeniero Sergio Matute, de centro de servicios de cómputo de la UNAM. En lo que respecta al software del sistema constituye una creación del ingeniero Sergio Matute¹⁵¹.

1.9.14. Colombia

Ha tenido grandes acercamientos a la Informática Jurídica, así en 1972 se inicia un proyecto para crear un banco de datos en legislación para el Congreso de la República, proyecto que se denominó SIDOC, el cual se suspende un año después y se retoma por el Banco Popular con el nombre SILCO, en donde se trabaja durante 4 años en el análisis e incorporación de información legal. Este proyecto fracasa por tener varias limitaciones en su diseño. Posteriormente se comienzan a crear bancos de datos circunscritos a áreas específicas; así la fundación Mariano Ospina Pérez crea una base de datos sobre Derecho Agrario, la Dirección de Impuestos Nacionales otra sobre Derecho Tributario, etc¹⁵².

¹⁵¹ Ibídem. p. 92

¹⁵² Ibídem. p. 93

En 1983 el Senado de la República quiso revivir la idea de un banco de datos en el que se incorporara toda la legislación nacional, por lo que el entonces ministro de Justicia, doctor Rodrigo Lara Bonilla, propició una serie de reuniones con la participación de la Oficina Latinoamericana de Informática Jurídica y Recursos Humanos de la Presidencia de la República, el Instituto Colombiano de Ciencia y Tecnología, Colciencias, y las distintas entidades y personas involucradas en estas actividades. Surge entonces la idea de celebrar un seminario para analizar las distintas orientaciones de la informática jurídica documental en el mundo. El seminario se realizó en 1984, contándose con la asesoría de la Corte de Casación de Italia, de Juris-Data de Francia y de la Asociación de Informática Jurídica y Derecho de España, en donde se concluyó que no era aconsejable para los países latinoamericanos comprometerse en la tarea de crear un banco jurídico global, sino que era más aconsejable definir una serie de pautas para la creación de bancos de datos, a las cuales se tuvieran que sujetar quienes fueran a crearlos, para garantizar la compatibilidad entre ellos.¹⁵³

Desafortunadamente para el país y para la Informática Jurídica muere el doctor Lara Bonilla, por lo que esta nueva ciencia en Colombia vuelve a tener un alto, siendo retomada posteriormente la idea por el Instituto Colombiano para el Fomento de la Educación Superior (ICFES), función que hoy en día se encuentra en cabeza de juriscol (dependencia adscrita al Banco de la República), quien hoy trabaja en la elaboración de bancos de datos jurídicos del ordenamiento jurídico colombiano. Actualmente, son muchas las instituciones que desarrollan bancos de datos en legislación, jurisprudencia y doctrina de las distintas áreas del saber jurídico.

¹⁵³ GIRALDO ÁNGEL, J. (1990). Op.cit., p.2.

- **Sijuec**

(Sistema de Informática Jurídica de la Universidad Externado de Colombia). Conscientes de la necesidad e importancia que para los abogados, jueces, magistrados, estudiantes, etc., tienen los fallos emitidos por las altas cortes para el desarrollo de sus correspondientes actividades, la Universidad Externado de Colombia, por conducto de su Departamento de Informática Jurídica elabora Bases de Datos que recogen las providencias judiciales emitidas por las altas cortes.

Aquí, se incorpora el texto completo de la providencia, y se hace un análisis señalando el problema o los problemas jurídicos en ellas contenidos, e indicando para cada uno de ellos las normas en que se basó el magistrado para resolverlo, el descriptor, como todos aquellos datos que identifiquen la providencia.

Los Bancos de Datos jurídicos, están grabados en un software bibliográfico denominado WINISIS. Adicionalmente la universidad desarrollo una interfase de consulta que hace fácil y útil la consulta para los usuarios¹⁵⁴.

En los tesauros temáticos, la información se debe recuperar a partir de los temas tratados en los respectivos documentos y no de las palabras utilizadas para presentarlos. Es precisamente esta la característica fundamental del sistema de Informática Jurídica de la Universidad Externado de Colombia (SIJUEC), la de ser un sistema diseñado para recuperar la información a partir del contenido conceptual

¹⁵⁴ VARGAS OSORNO T.G. ob. cit. p. 94.

de los documentos jurídicos, y no de las palabras que en él están escritos. Para ello fue necesario elaborar toda una metodología encaminada a establecer los principios que deben seguirse para elaborar tesauros temáticos, tomando muchos postulados de las técnicas bibliotecológicas, pero definiendo características particulares que son propias de la documentación de contenido jurídico.

En cumplimiento de su deber de mantenerse actualizada e ir a la vanguardia académica, la Universidad Externado de Colombia, merced a un trabajo inmenso y esmerado, incorpora al patrimonio cultural de la profesión jurídica la técnica y la metodología, siendo pionera en el ofrecimiento de los bancos de datos jurídicos: abogados, litigantes, jueces, fiscales, magistrados, empresas, instituciones, asesores, alumnos, profesores, investigadores, académicos e intelectuales podrán acceder al volumen íntegro y actualizado de la doctrina de las altas Cortes, así como de la legislación vigente concordada, en el disco compacto. En el pasado quedó la penosa consulta manual de la información jurídica, desde la creación de los bancos de datos de la Universidad Externado se podrá encontrar articulada y clasificada.

Fue muy satisfactorio para la Universidad Externado de Colombia - Departamento de derecho informático, presentar a estudiosos del derecho nueve primeros ejemplares de la colección de bancos de datos jurídicos. El primer ejemplar fue el de “Contratación Estatal”, que es una recopilación completa de legislación concordada y providencias de la Corte Constitucional, en la materia, desde su inicio hasta la fecha y del Consejo de Estado desde 1983 hasta la fecha, cuidadosamente analizadas y clasificadas de acuerdo con la metodología jurídica y los tesauros.

El segundo ejemplar fue el de “Derecho Penal”, integrado por la legislación vigente concordada y las providencias proferidas por la Sala de Casación Penal de la Corte Suprema de Justicia desde 1982 hasta la fecha, y Corte Constitucional, en la materia desde su inicio hasta la fecha, cuidadosamente analizadas y clasificadas de acuerdo con la metodología jurídica y el tesauros , en las que se plantean y solucionan los distintos problemas jurídicos que giran alrededor del desarrollo y perfeccionamiento teórico práctico del derecho penal en general, como mecanismo de control social formal que integra el sistema penal.

El tercer ejemplar fue el de “Derecho Laboral y Seguridad Social”, integrado por la legislación vigente concordada y las providencias proferidas por la Sala de Casación Laboral de la Corte Suprema de Justicia, Consejo de Estado, Sala de lo Contencioso Administrativo, Sección Segunda desde el año 2000 hasta la fecha y Corte Constitucional, en las que se plantean y solucionan los distintos problemas jurídicos que giran alrededor del desarrollo y perfeccionamiento teórico-práctico del derecho laboral.

El cuarto ejemplar fue el de ”Derecho de Familia e Infancia” integrado por la legislación vigente concordada y las providencias proferidas por la Sala de casación civil de la Corte Suprema de Justicia desde 1973 hasta la fecha y de la Corte Constitucional, en la materia, desde su inicio hasta la fecha, cuidadosamente analizadas y clasificadas de acuerdo con la metodología jurídica y el tesauo, en las que se plantean y solucionan los distintos problemas jurídicos que giran en torno del Derecho de Familia.

El quinto ejemplar fue el de “Derecho de Seguros” integrado por la legislación vigente concordada y las providencias proferidas por la Sala de casación civil de la Corte Suprema de Justicia, Consejo de Estado y Corte Constitucional, desde 1970 hasta la fecha cuidadosamente analizadas y clasificadas de acuerdo con la metodología jurídica y el tesoro, en las que se plantean y solucionan los distintos problemas jurídicos que giran en torno del Derecho de Seguros.

El sexto ejemplar fue el de “Derecho Privado” (obligaciones y contratos) que es una recopilación completa de legislación concordada y providencias emitidas por la Sala de Casación Civil de la Corte Suprema de Justicia y Corte Constitucional, desde 1918 hasta la fecha, cuidadosamente analizadas y clasificadas de acuerdo con la metodología jurídica y el tesoro.

El séptimo ejemplar fue el de “Derecho Tributario” que es una recopilación completa de legislación vigente concordada y providencias emitidas por el Consejo de Estado, Jurisdicción de lo Contencioso Administrativo desde el año 2000 hasta la fecha, y de las providencias emitidas por la Corte Constitucional, en la materia, desde su inicio hasta la fecha. Conceptos de la DIAN desde el año 2004 hasta la fecha, cuidadosamente analizadas y clasificadas de acuerdo con la metodología jurídica y el tesoro.

El octavo ejemplar fue el de “Derecho Aduanero y Cambiario” que es una recopilación completa de legislación vigente concordada y providencias emitidas por el Consejo de Estado, Sección Primera, desde el año 2000 hasta la fecha y las

emitidas por la Corte Constitucional, en la materia, desde su inicio hasta la fecha.
Conceptos de la DIAN desde el año 2000 hasta la fecha.

Se creó además, el banco de datos de “Derecho del Medio Ambiente”, que es una recopilación completa de legislación vigente concordada y providencias de la Corte Constitucional, en la materia, desde su inicio hasta la fecha y del Consejo de Estado, Sala plena y Sección Tercera, desde 1974 hasta la fecha.

En el año 2011, surge en esta misma Casa de Estudios, el banco de datos jurídico de “Acciones Populares y de Grupo”, que recopila la totalidad de la materia desde sus inicios, ella comprende la legislación y jurisprudencia de las altas Cortes con un análisis sobre cada uno de los documentos jurídicos realizado por un especialista. Es importante mencionar, que en la actualidad, se encuentra en desarrollo el banco de datos de “Derecho Informático” que pretende presentar la primera ontología aplicada al derecho.

En la elaboración de dichos Bancos de Datos se resaltan como puntos principales el desarrollo de **TESAUROS JURIDICOS** para cada una de las áreas, la metodología jurídica para el análisis de la jurisprudencia (problema jurídico, tesis, fuente formal, descriptores y texto completo).

Dentro del análisis de providencias judiciales señalamos:

1°. Datos que identifican el documento: tales como el autor, sea corporativo o personal, título de la providencia que indica la fecha de la misma, nombre del

Magistrado o Consejero Ponente, demandante, demandado, expediente, notas en donde se incluyen las sentencias y autos en igual sentido y en sentido contrario que ha señalado la misma providencia que se analiza.

2°. Análisis: al estar frente a una providencia nos planteamos el problema o problemas jurídicos a los cuales la Corporación les está dando respuesta. A cada problema jurídico se le extrae la respuesta que constituye la tesis que se incorpora a la base de datos. Esta tesis es textual.

También a cada problema jurídico se le señala su correspondiente texto. Al problema jurídico se le asigna la fuente jurídica que la Corporación tuvo en cuenta para darle solución al mismo.

Todo esto se titula con un descriptor o varios sacados de los tesauros jurídicos debidamente elaborados en cada área por el departamento.

En cuanto a su consulta, las bases de datos se crearon para que la búsqueda fuera sencilla y con el fin de que el usuario tenga la posibilidad de hacer distintas combinaciones sin tener un conocimiento profundo sobre computadores. Se utiliza el software bibliográfico WINISIS mantenido y diseminado por UNESCO, para registrar los datos identificadores de cada una de las providencias que van al Banco de Datos.

Se utiliza el procesador de textos para la grabación de los textos. La Universidad desarrolló un programa interface que hace amigable la consulta y permite la unión de

lo grabado en WINISIS con el texto o extracto grabado en el procesador de palabra, es así como el usuario final puede consultar los bancos de datos en un CD ROM.

En este orden de ideas, las opciones para consulta de cada una de las bases de datos son:

POR TEMA: Por los descriptores asignados de acuerdo con el tesauro elaborado para cada área.

POR TITULO: El cual está conformado por la denominación del documento (ley, decreto, resolución, sentencia, auto, salvamento de voto, etc.) y la fecha.

POR FECHA: Fecha del documento.

POR AUTOR PERSONAL: Permite la consulta de los salvamentos y aclaraciones de voto hechos por un magistrado.

POR PONENTE: Índice que contiene los nombres de los magistrados de la respectiva Sala.

POR FUENTE FORMAL: Índice que contiene las fuentes jurídicas que sustentan cada una de las tesis de la Corporación en sus respectivas áreas.

POR DEMANDANTE: Índice que contiene el listado de demandantes.

POR DEMANDADO: Índice que contiene el listado de demandados.

POR EXPEDIENTE: Índice que contiene el número de radicación de las providencias.

La consulta puede realizarse independientemente por cada uno de los campos señalados anteriormente o mediante la combinación con los conectores lógicos **Y** u **O**. El usuario tiene la opción de ver e imprimir lo consultado que consta de los datos identificadores de la providencia, tesis y texto. Además puede enviar a su procesador de palabra lo que desee después de hacer la consulta, como también buscar palabras dentro del texto.

2. BANCOS DE DATOS JURÍDICOS: PRESENTE

2.1. Introducción

Estudiamos dentro del acápite anterior, el surgimiento de la informática jurídica, sus conceptos fundamentales, las primeras formas de creación, almacenamiento, organización y clasificación de la información y sus más importantes problemas, mencionábamos que, uno de los inconvenientes que presentaba era la capacidad de almacenamiento de las bases de datos, lo anterior, a causa del espacio reducido, con el cual contaban los ordenadores y medios de almacenamiento informáticos de la época.

En el interregno entre la informática jurídica documental del pasado y la del presente, el surgimiento de Internet, la informática jurídica documental, que hasta el momento había sido la más desarrollada tuvo un alto en el camino. Los estudiosos de ella no se volvieron a preocupar por muchas de las cosas que antes eran de preocupación constante, ya que el avance de la tecnología solucionó gran parte de su objeto de estudio, como era el espacio en disco duro, que en su momento se vieron obligados a no ingresar el texto completo sino a incluir extractos.

Una vez examinado el pasado de la informática jurídica documental, podemos llegar a la conclusión que las personas que se interesaron por el estudio de esta ciencia, centraron toda su atención, en los problemas que se presentaron en el tratamiento de la información y de los instrumentos lingüísticos para la solución de los mismos.

No obstante, como expondremos en lo sucesivo de esta segunda parte, el problema anteriormente planteado ha sido solucionado gracias a la creación de Internet y su desarrollo mediante World Wide Web, Web 1.0 y Web 2.0 que en la actualidad cuenta con un total de 19.55 billones de páginas, indexadas en el contenido de la Web por los buscadores Google, Bing, Yahoo y Ask¹⁵⁵, de forma que se superó satisfactoriamente el problema anteriormente planteado de la capacidad de almacenamiento de información en la bases de datos.

Sin embargo, teniendo en cuenta que la capacidad de almacenamiento se ha incrementado ostensiblemente, pese a que desapareció el problema que se presentaba en el pasado, esta solución, trajo consigo la inflación de información, también denominada “hiperinflación,”¹⁵⁶ que se caracteriza por un alto contenido de información documental que reposa en internet, dentro de la cual los usuarios que emplean los buscadores como herramientas de selección de información, se encuentran las más de las veces con un alto contenido de información dentro del resultado de la búsqueda, sobre la cual el usuario deberá desarrollar la actividad de discriminación y verificación de la confiabilidad de la misma, pudiendo caer dentro del curso de este trabajo, en errores que pueden llegar a alterar los resultados reales, oportunos, pertinentes y veraces esperados por el usuario.¹⁵⁷

¹⁵⁵ DE KUNDER, M. (2011, Julio 12). *The size of the World Wide Web. The Internet.* (U. d. Tilburg, Ed.) <http://www.worldwidewebsite.com/>

¹⁵⁶ “Cada vez es más fácil encontrar en Internet documentación de toda clase referida a los países del Primer Mundo (y muy especialmente, a Estados Unidos), mientras que hay grandes dificultades para documentarse sobre temas vinculados a los otros países”. SIRI, L. (2000). *Internet: búsquedas y buscadores.* (Vol. 2). Bogotá: Norma. p. 15.

¹⁵⁷ Ibídem. p. 19. “La presencia en la Web de una enorme cantidad de material irrelevante, que convierten las búsquedas en el hallazgo de una aguja en un pajar”.

De esta forma, pese a que el usuario tenga los conocimientos técnicos necesarios este puede ser víctima de la desinformación a causa de la inflación de información, incluso en los bancos de datos especializados. “El problema real es que es demasiado fácil encontrar la basura y demasiado difícil encontrar lo que se necesita. Y, a diferencia de otros aspectos de la Internet, este problema afecta a todos, desde novatos a expertos, desde usuarios a desarrolladores y desde gente de negocios hasta maestros de escuela”¹⁵⁸.

El problema más grave de la inflación de información, es que puede llegar en múltiples ocasiones a suministrar información errada, por esta razón la pregunta que nos debemos plantear es: “Quien está en condiciones idóneas de ejercer esa función mediadora, similar a la que debieran ejercer los periodistas entre las fuentes y el público, en Internet: un medio caótico, sobre dimensionado, con autoridad desdibujada y donde (a diferencia de la televisión o los diarios) cualquiera puede en principio publicar cualquier cosa”¹⁵⁹.

De otra parte, existen otros problemas que desafortunadamente continúan subsistiendo pese a la existencia de Internet, estos son: Recuperación de la información, organización de la búsqueda de información y clasificación de la información. Motivo por el cual nos ocuparemos dentro del presente capítulo, de analizar que herramientas ofreció desde sus inicios internet, cuales nos ofrece en la actualidad, que ventajas y desventajas trajo consigo el surgimiento de los nuevos buscadores generales, especializados y sus repercusiones en la informática jurídica.

¹⁵⁸ Ibídem. p. 29.

¹⁵⁹ Ibídem. p. 38.

En consecuencia, analizaremos el estado actual de las bases de datos, buscadores o search engine, para estos efectos, será necesario, en primer lugar analizar los antecedentes de internet, el surgimiento del World Wide Web, los conceptos básicos de web 1.0, HTML, hipervínculo, lenguaje de marcación, web 2.0, el formato XML y su incidencia en las bases de datos jurídicas que existen en la actualidad.

En segundo lugar, será necesario examinar exhaustivamente los buscadores generales, aquellos especializados y su forma de organizar la información. Para finalmente identificar y analizar los bancos de datos jurídicos que continúan del pasado y los que han surgido de forma posterior en la web, a los cuales tenemos acceso en la actualidad y confrontar dicha información a la luz de la informática jurídica. Con el fin de poder establecer si las herramientas suministradas en la actualidad por los buscadores que ofrece internet son útiles o no para efectos de buscar información jurídica y en particular reconocer la vital importancia de los bancos de datos jurídicos especializados en todo el mundo.

2.2. Historia de Internet

Antes de iniciar el estudio de la historia de internet, es imprescindible realizar una explicación sobre su etimología y concepto.

La palabra internet proviene del inglés *internet* formado de *inter* (entre) y *net* de network (red electrónica)¹⁶⁰. Internet está definido como un conjunto de redes de

¹⁶⁰ DECHILE.NET. (2001). Consultada en enero 17, 2015. <http://etimologias.dechile.net/?Internet>

comunicación con alcance mundial, descentralizado que permite la conexión de computadores, el intercambio y acceso a la información desde cualquier lugar.

Igualmente, Internet ha sido definido como “una red global en la que se conjugan todas las redes que utilizan protocolos TCP/IP y que son compatibles entre sí. La red Internet se compone de una cantidad ilimitada de ordenadores interconectados entre sí”¹⁶¹.

El surgimiento de internet, se remonta al siglo XX, cuando se crea un protocolo de control de transmisiones, que permite el trabajo de los computadores en red desde diferentes lugares, hecho que cambia de forma contundente el curso de la informática jurídica como lo veremos más adelante.

Tenemos pues, que dentro del afán de la búsqueda de conexión para la comunicación entre computadoras, surge *arpanet*¹⁶², que es catalogada como el origen prístino de Internet. Esto fue formulado por primera vez en 1962, por Joseph Carl Robnett Licklider de la empresa *Bolt, Beranek and Newman* (BBN) compañía de alta tecnología, que discutía sobre la creación de una red galáctica.

En octubre de 1962, Joseph Carl Robnett Licklider es convocado a participar en un foro por ARPA (Agencia de Investigación de Proyectos Avanzados) perteneciente al Departamento de Defensa de los Estados Unidos. No obstante mal haríamos en

¹⁶¹ FERNÁNDEZ-COCA, A. (1998). *Producción y diseño gráfico para la World Wide Web*. Barcelona: Paidós Ibérica. p. 35.

¹⁶² “Arpanet en sus siglas *Advanced Research Projects Agency Network*, fue creada por encargo de Departamento de Defensa de los Estados Unidos, como medio de comunicación para los diferentes organismos del país. Después ARPA se transforma en DARPA para la adjunción de D de defensa”. BÉRA, M., & MECHOULAN, É. (1999). *La machine internet*. Paris: Ediciones Odile Jacob. p. 18.

atribuirle la creación de Internet exclusivamente a Licklider, por lo que es importante mencionar que el origen de Internet también contó con la participación y el aporte de tres expertos en tecnología: Paul Baran, Leonard Kleinrock y Donal Davies.

El primero de ellos, *Paul Baran*, perteneciente a la Corporación RAND Investigación y desarrollo, (Research And Development) por sus siglas en inglés, entidad formada en principio con el fin de apoyar a las fuerzas armadas norteamericanas con investigación y análisis, en 1959 inicia una investigación, debido a su afán por crear una red segura de comunicaciones que pudiese sobrevivir a ataques nucleares, resultados que se publicaron en el año de 1969 y aportaron dos ideas importantes para la creación de internet: La utilización de una red descentralizada, que se encuentre capacitada para solucionar sus propias fallas y la división de mensajes completos en fragmentos que sigan caminos distintos.

En segundo lugar, encontramos a *Leonard Kleinrock*, quien adelantaba su tesis doctoral en el Massachusetts Institute of Technology (MIT), donde trabajó el concepto de almacenamiento y reenvío de mensajes, estudio que incluyó un importante aporte a las redes de comunicaciones, que tuvo al igual que el aporte de Paul Baran y Licklider una incidencia considerable en el origen de internet.

El último de los expertos en tecnología fue Donal Davies, quien desde el *Laboratorio Nacional de Física*, relacionó los conceptos anteriores a excepción del trabajo de Baran sobre el cual tuvo conocimiento después de realizar su investigación. Así pues, vemos cómo influyen de forma independiente estos cuatro centros de investigación en la creación de internet, centros que dicho sea de paso, en

la conclusión de los estudios terminaron convirtiéndose en los primeros nodos experimentales de *Arpanet*.

Los primeros computadores de *Arpanet* en conectarse, fueron el SDS Sigma 7 y un IBM 360 en la Universidad de California, en Los Ángeles, donde *Leonard Kleinrock*, creó el Centro de medición de red, a su vez en el Instituto de Investigación de Standford, mediante un sistema de hipertexto, conectó un SDS 940, igualmente, Iván Sutherland conectó en el departamento gráfico de la *Universidad de Utah* un PDP-10.

Con los computadores antes mencionados, se realizó el primer enlace el 21 de noviembre de 1969 entre la Universidad de California en Los Ángeles y la de Standford localizada en la misma ciudad. Posteriormente los cuatro computadores iniciales ya formaban una red para el 5 de diciembre de 1969. Al año siguiente esta red inicial llegó a la costa Este donde se unió otro ordenador de la BBN, en 1971 ya formaban parte de la red 23 computadores pertenecientes a centros de investigación de las diferentes universidades, el número fue creciendo; para el año de 1981, ya existían 213 computadores conectados a la red inicial de *Arpanet*, finalmente, para el año de 1983 ya se contaba con 500 computadores.

De manera simultánea a lo anterior, Ray Tomlinson creó el correo electrónico, suceso que contribuyó de forma ostensible al incremento del uso de Internet¹⁶³, igualmente, otro de los servicios que aumentó el uso de Internet de forma considerable, fue la creación de la World Wide Web (WWW), que fue desarrollado

¹⁶³ CALVO PRIETO, J. M., ROCHA FREIRE, I., DESONGLES CORRALES, J., & SANTOS RODRÍGUEZ, M. (2006). Auxiliar de servicio de la Universidad Politécnica de Valencia (Vol. 2). España: Editorial Mad S.L. p. 164.

después del año 1990 y facilitó la consulta de archivos de hipertexto. Cabe señalar en este punto, que el concepto de web e Internet se suelen confundir de forma frecuente, no obstante, la clara diferencia reposa en que Internet es el medio de transmisión que utiliza la World Wide Web.¹⁶⁴

Posteriormente, la primera red “*Arpanet*”, siendo la única red existente, que además prescindía del control de errores en los host, evolucionó hacia el Internet actual, creando redes independientes y abiertas con diferentes host, incluyéndose la comunicación mediante redes de paquetes por satélite. Estas redes independientes, se conectaban de computador a computador, pretermitiendo absolutamente las características particulares de cada uno de ellos. La creación, el diseño y elaboración del proyecto se le atribuye a Robert Kahn.

El mencionado proyecto, radicado en cabeza de Kahn, posteriormente se denominó *Transmission Control Protocol* (TCP), que principalmente buscó remplazar el protocolo usado en *Arpanet*. El diseño del TCP, contenía las siguientes características:

“Pequeñas sub-secciones de la red serían capaces de hablar entre sí mediante un computador especializado que simplemente reenvía paquetes (inicialmente llamado gateway, hoy en día conocido como router).”

¹⁶⁴ La puesta en marcha del World Wide Web, únicamente fue posible, gracias a la implementación del servicio TCP y de IP, como veremos en líneas posteriores dentro de las cuales se explicaran estos dos protocolos.

Ninguna porción de la red podría constituir un punto de ruptura (de forma que toda la red dependiese de su correcto funcionamiento), ni sería capaz de tomar control sobre toda la red.

Cada trozo de información enviado a través de la red recibiría un número de secuencia, para asegurar que fuese procesado en el orden adecuado a su llegada al destino. Este número sería también usado para detectar la pérdida de información en tránsito.

*Un computador que envíe información a otro debería saber si ésta ha sido recibida cuando el destinatario devuelva un paquete especial, conocido como *acknowledgement* (confirmación, acuse de recibo), para cada trozo de información enviado.*

Si la información enviada se perdiera, sería retransmitida una vez se haya superado el debido tiempo de espera. Cada trozo de información enviado a través de la red iría acompañado de un checksum¹⁶⁵, calculado por el emisor original, y comprobado por el receptor, para asegurar la ausencia de daños durante el transporte”¹⁶⁶.

¹⁶⁵ El checksum, o suma de verificación, es una forma de medida para proteger la integridad de los datos. Además de ser usado para internet, es utilizado para la comunicación de dispositivos y datos almacenados. Estas medidas se realizan sumando los componentes del sistema para que una vez almacenado este valor, se pueda comparar el resultado con el valor almacenado y así finalmente poder comprobar si las sumas concuerdan o no, lo que determinará si ha sido corrompido el sistema.

¹⁶⁶ Robert Kahn es.wikipedia.org. (2015). Consultada en agosto 28, 2014.
https://es.wikipedia.org/wiki/Robert_Kahn

Seguidamente, en 1973 Vint Cerf de la Universidad de Stanford se unió con Robert Kahn con el fin de terminar la primera versión de TCP, finalmente esta versión inicial se separó en dos capas, desplazando las funciones más básicas al *Internet Protocol* (IP). Hoy por hoy son conocidas como TCP/IP y es la base del internet del siglo XXI.

2.3. La Incidencia de la World Wide Web en las Bases De Datos

Antes de hablar sobre la incidencia de la world wide web en las bases de datos, es necesario mencionar que contrario a lo que se podría pensar, www guarda absoluta diferencia con Internet a pesar de su siempre pretendida asimilación, es así como se ha sostenido que “la World Wide Web es como una enciclopedia, una guía telefónica, una colección de discos, una tienda de video, altavoces y de la esquina, todo en uno y accesible a través de cualquier ordenador. Se ha vuelto tan exitoso que para muchos es sinónimo de Internet, pero en realidad los dos son muy diferentes. Internet es como una red de caminos electrónicos que cruzan el planeta, en la muy promocionada superautopista de la información. La web es sólo uno de los muchos servicios que utilizan la red, al igual que muchos tipos diferentes de vehículos de uso de las carreteras. En Internet, la web resulta ser de lejos la más popular”.¹⁶⁷

¹⁶⁷ “The World Wide Web is like an encyclopaedia, a telephone directory, a record collection, a video shop, and speakers corner all rolled into one and accesible through any computer. It has become so successful that to many it is synonymous with the Internet, but in reality the two are quite different. The internet is like a network of electronic roads criss-crossing the planet -the much- hyped information superhighway. The web is just one of many services using that network, just as many different kinds of vehicle use the roads.

On the Internet, the web just happens to be by far the most popular. The arrival of the web in 1990 was to the Internet like the arrival of the internal combustion engine to the country lane. Internet transport would never be the same again”. GILLIES, J., & CAILLIAU, R. (2000). *How the Web was Born: The Story of the World Wide Web*. New York: Oxford University Press. p. 1.

“La web comenzó en el laboratorio europeo de partículas físicas y energía conocido como CERN. En 1989, el físico Tim Bernes propuso el concepto de web como un sistema para transferir ideas e investigaciones entre la comunidad de científicos relacionados con la física y la energía. La comunicación era decisiva para este grupo localizados en todo el mundo.

La propuesta definía un sistema simple que usaría hipertexto: Una forma de presentar y relacionar información con enlaces en lugar de líneas secuenciales. Al principio, el programa no permitía transmitir imágenes, audio ni video”¹⁶⁸. Cabe señalar, que la World Wide Web, ha sido definida como “el ambiente gráfico a través del cual usted puede ingresar a una inmensa selección de datos en forma de textos, gráficos, dibujos, sonidos y video”¹⁶⁹.

En este orden de ideas, se define a la World Wide Web o en su traducción literal: “telaraña mundial”, como aquel sistema que permite la difusión de la información por conducto de Internet que hace las veces de transportador, de forma que bien se puede considerar como un medio de comunicación dentro del cual se alojan contenidos de imágenes, texto, gráficos y otras herramientas multimedia.

La web cuenta con un sistema de hipertexto¹⁷⁰ HTTP (Hypertext Transfer Protocol o Protocolo de Transferencia de Hipertexto) cuya función principal consiste en realizar la totalidad de transacciones en la www, en otras palabras, el sistema HTTP es el método con el cual se transfieren las páginas web al computador del usuario.

¹⁶⁸ ARANGO VÁSQUEZ, S. I., & RICAURTE AVENDAÑO, A. (2006). *Manual de herramientas tecnológicas*. Medellín: Universidad de Medellín. p. 46.

¹⁶⁹ *Ibíd.*

¹⁷⁰ El hipertexto es el nombre que se le otorga a aquellos documentos que reposan en un dispositivo electrónico, contentivos de opciones que direccionan hacia otros contenidos.

Finalmente es importante dejar claro que la World Wide Web no es más que una de las aplicaciones que se encuentran en Internet. Adicionalmente, existen otras como el correo electrónico o la mensajería instantánea.

El Protocolo de Transferencia de Hipertexto HTTP, “es el protocolo utilizado para el intercambio de datos entre el servidor y los clientes de la www. Se podría interpretar como una serie de simples mensajes, que puede ser guiado de manera asincrónica entre encaminadores (TCP/IP), que describen enteramente la función a realizar en el ordenador remoto e indican quien formula la solicitud. Al llegar el mensaje es simplemente analizado y ejecutado”¹⁷¹.

Con el surgimiento de la World Wide Web, notoriamente se presentó el cambio inusitado más contundente y determinante para las bases de datos antiguas, este no fue otro que la creación de la más grande y completa biblioteca existente en el mundo. Uno de los primeros populares navegadores web, modelado después de Hyper Card, fue Viola WWW.

No obstante el punto decisivo para la World Wide Web comenzó con la introducción de Mosaic en 1993, un navegador Web con interfaz gráfica desarrollado por un equipo en el National Center for Supercomputing Applications en la Universidad de Illinois, esta interfaz gráfica, pronto se hizo muy popular, dado que la interfaz gopher, estaba basada en texto de forma que la World Wide Web se convirtió en la interfaz preferida para acceder a Internet.

¹⁷¹ CAFASSI, E. (1998: P). *Internet olíticas y comunicación*. Buenos Aires, Argentina: Biblos. p. 260.

En 1994, Netscape Navigator de Andressen, se posiciona como el navegador web más famoso en el mundo, remplazando a Mosaic. Por otra parte la creación de Internet explorer y una amplia variedad de navegadores, han sustituido drásticamente esta interfaz.

“El éxito de Internet y, en concreto, de la World Wide Web, ha desencadenado un cambio tecnológico trascendental en la industria de la información; la Web, convertida en interfaz única, ha alterado radicalmente el mercado de los contenidos en cuanto a la producción, almacenamiento, difusión y recuperación de la información y, todo ello, a una velocidad vertiginosa. En la segunda mitad de la década de los noventa, la Web pasa a convertirse en plataforma óptima para la distribución de contenidos jurídicos. Así, desde 1996 los diversos editores jurídicos comienzan paulatinamente a disponer de sedes web y se inicia la inevitable migración de las bases de datos al nuevo entorno.”¹⁷².

Teniendo en cuenta lo anterior, nos corresponderá analizar en otros acápites posteriores, como surgieron los buscadores y bancos de datos capaces de seleccionar, discriminar y organizar la gran cantidad de información que se encuentra en Internet, de forma tal, que los usuarios acuden en la actualidad en un alto porcentaje a consultar los motores de búsqueda o *search engines* diariamente en la Web,¹⁷³ con el

¹⁷² ALVITE DÍEZ, M. L. (2009). op.cit. p. 2.

¹⁷³ “84% de los estadounidenses en línea han utilizado motores de búsqueda - que se traduce en más de 107 millones de personas. En cualquier día determinado en línea, más de la mitad de los que utilizan los motores de búsqueda el uso de Internet. Y más de dos tercios de los usuarios de Internet dicen que usan los motores de búsqueda al menos un par de veces por semana.

El uso de motores de búsqueda por lo general ocupa el segundo lugar a utilizar el correo electrónico como la actividad más popular en línea. Durante los períodos en que las historias importantes de

único fin de poder localizar la información alojada dentro de ese gran universo que reposa en la Web y que se encuentra en constante movimiento.

Otro punto neural, que resulta necesario mencionar dentro de este acápite, es la organización de la World Wide Web, que principalmente cuenta con dos sistemas para almacenar y recuperar la información en ella contenida, estos son:

noticias se están rompiendo, el acto de puesta en línea de noticias por lo general supera el uso de motores de búsqueda.

Hay un pago sustancial como motores de la búsqueda de mejorar y las personas se vuelven más hábiles en el uso de ellos. Alrededor del 87% de los usuarios del motor de búsqueda dicen que encontrar la información que desean la mayoría de las veces cuando se utilizan los motores de búsqueda.

La conveniencia y la eficacia de la experiencia de búsqueda solidifica su recurso de casación. Aproximadamente el 44% dice que la mayoría de veces que la búsqueda que están buscando información vital que tienen una necesidad imperiosa.

Estadounidenses realizaron 3.900.000.000 totales de búsquedas en junio.

Proyecto de investigación Pew Internet en los últimos años ha demostrado consistentemente que el uso de motores de búsqueda es una de las actividades online más populares. Durante la mayoría de los meses, el acto de usar un motor de búsqueda sólo es superado por correo electrónico como la actividad de Internet más frecuentes. (Durante los períodos en que las historias importantes de noticias se producen, el acto de puesta en línea de noticias supera los motores de búsqueda por un pequeño margen. Pero no hay nada en nuestra investigación ha llegado cerca de igualar la popularidad del correo electrónico.)

El 84% de los usuarios de Internet, más de 107 millones de estadounidenses, el informe que han utilizado motores de búsqueda. El volumen de búsquedas es impresionante. datos de comScore muestran que un usuario medio de Internet realiza 33 búsquedas por mes, sumando así más de 3,9 millones de búsquedas por mes en total 25 motores de búsqueda más populares de Internet. Por otra parte, la frecuencia de la búsqueda es impresionante, también.

Más de un tercio de los buscadores (35%) reportan que iniciar una búsqueda por lo menos una vez al día, (con dos tercios de ese grupo de información que utilizan los motores de búsqueda varias veces al día.) Otra tercera búsqueda varias veces durante la semana (media de la búsqueda 1 - 2 días a la semana y media de búsqueda 3-5 días a la semana). El último tercio de la búsqueda con menos frecuencia que eso.

Preguntado acerca de sus experiencias y sentimientos utilizando los motores de búsqueda, la mayoría de los estadounidenses expresan su satisfacción, confianza y confianza en sí mismo. Alrededor del 87% de los usuarios del motor de búsqueda dicen que encontrar la información que están buscando la mayoría de las veces realizar búsquedas. Alrededor del 20% de ese grupo dicen que siempre tienen éxito. Más de dos tercios de los usuarios del motor de búsqueda (68%) dicen que consideran los motores de búsqueda de una fuente imparcial y objetiva de la información. Y los usuarios de búsqueda de sí mismos aplaudir también. Un increíble 92% de los usuarios expresan la confianza en sus habilidades como investigadores - más de la mitad de ellos dicen que están "muy confiados" de que puede lograr lo que quieren cuando realizan una búsqueda en línea.

FALLOWS, D., LEE, R., & GRAHAM, M. (2004, Agosto 12). *Memo de datos en los motores de búsqueda*.

Consultada en febrero 5, 2014. Pw Research Center:

<http://translate.google.com/translate?hl=es&langpair=en%7Ces&u=http://www.pewinternet.org/Reports/2004/Data-Memo-on-Search-Engines/Data-Memo.aspx>.

“El desarrollo de search engines, o robots que buscan contenidos de la web por keywords, o palabras clave. Son herramientas muy útiles, aunque tienden a devolver gran cantidad de información irrelevante (entre otros problemas).

El Otro enfoque consiste en producir listados temáticos que alienten a los usuarios a navegar por la web. Estos listados jerárquicos tienden a adoptar esquemas clasificatorios bibliotecológicos”¹⁷⁴.

En la actualidad, encontramos que los *search engines*, cuentan con mayor acogida entre los consumidores, sin embargo, esto no quiere decir que el sistema para encontrar recursos de Internet de búsqueda mediante *keywords* sea aquel que contiene la información precisa y completa que el usuario pretende encontrar en la Web.

Por lo anterior, han sido varias las recomendaciones que al respecto se han hecho sobre el uso de los *listados temáticos*, como medio idóneo para encontrar información clasificada de forma correcta en términos de bibliotecología. Por lo cual, se hace necesario mencionar sucintamente que los listados temáticos tienen varias ventajas dentro de las cuales podemos encontrar que los temas se encuentran clasificados de forma que ofrecen justamente el contexto del objeto de la búsqueda, por otro lado, al ser una clasificación por consenso ofrece la posibilidad de realizar búsquedas entre diferentes bases de datos.

¹⁷⁴ SIRI, L. (2000). op.cit. p. 51.

Mencionado lo anterior, podríamos llegar a afirmar que la informática jurídica documental, dentro de estas etapas presenta los mismos planteamientos que los buscadores: En un primer momento, se presenta el problema de la capacidad de almacenamiento, la inclusión parcializada de los documentos jurídicos, que conllevó a afectar el procedimiento de análisis subjetivo del usuario, que se veía obligado a sujetarse al albedrío de la persona encargada de llevar la información al banco de datos, que las más de las veces era un jurista, quien subjetivamente determinaba que parte del documento iba a encontrarse contenido en el mismo.

Con el surgimiento de la world wide web y en particular de la creación de los servidores o Hostings se resolvió el problema de almacenamiento que se presentaba en el pasado, dado que es un recurso que proporcionó el cambio del concepto de almacenamiento de información y en el caso que nos interesa de la informática jurídica, otorgó nuevas herramientas para alojar la totalidad de los documentos jurídicos, toda vez que como bien se puede recordar, uno de los problemas más sobresalientes era precisamente el espacio de almacenamiento de información que conducía en varias ocasiones a resumir el contenido del documento privando al usuario de la posibilidad de la consulta íntegra del mismo.

El servidor o hosting, “es el servicio de puesta en línea de material de Internet, por ejemplo una página web”¹⁷⁵, en nuestro concepto el hosting es aquel lugar donde reposa la totalidad de la información de una página web.

¹⁷⁵ CAFASSI, E. op. cit. p. 260.

De otro lado, en un segundo momento, existe un planteamiento que aún persiste y que corresponde a los problemas de localización de información actualizada, pertinente, oportuna, toda vez que como veremos a continuación la world wide web es ocupada por los incalculables contenidos de la web 1.0, la web 2.0 y porque no decirlo, por la web 3.0 aunque en una menor medida.

2.4. Web 1.0

Es la forma originaria en la cual se empezaron a presentar las páginas web, esta ofrece información para consulta a los usuarios, en una sola dirección que va desde el web master o creador de la página al usuario, de manera que su interacción con los contenidos es nula, en este sentido, la restricción de creación de contenidos era ostensible. La web 1.0 ha sido y continúa siendo la abanderada a la hora de hablar de páginas web que cuentan con bancos de datos jurídicos en la actualidad.

La definición de la web 1.0, no hace referencia a un avance en la tecnología de la web, toda vez que como estudiaremos en el acápite siguiente la web 2.0 cuenta con una técnica diferente para el diseño de la página web y en particular para contribuir con sus contenidos y generar interacción con los usuarios. Por lo anterior, es imposible definir la web 1.0 en un periodo determinado diferente al de la web 2.0 toda vez que en la actualidad son coexistentes, ninguna de las dos excluye a la otra y la web 1.0 hace referencia a un estado de la tecnología de presentación de las páginas web.

Como características de la web 1.0, podemos mencionar varias, entre ellas que: Son sitios webs estáticos, en el sentido que proporcionan información pero no existen motivos para volver a consultarlas toda vez que sus contenidos serán siempre los mismos, tal es el caso de las páginas web de oficinas de abogados que ofrecen sus servicios, la gran mayoría de las veces estas se crean con el fin de proporcionar al usuario información de contacto. Otra característica, consiste en que la tecnología asociada a la web 1.0 es HTML y GIF, que son los formatos con los cuales se presentan las páginas web en la web 1.0 de forma que el contenido de las mismas se crea de una forma fija y el usuario que accede a esta información siempre encontrará los mismos contenidos, en otras palabras es totalmente estática.

Otras características que identifican a la web 1.0, consisten en el uso de elementos de diseño tales como los framesets o marcos,¹⁷⁶ los libros de visitas on line o guestbooks¹⁷⁷, al igual que los botones GIF o botones animados para las páginas web y etiquetas que son extensiones del HTML, entre otros.

En cuanto a la interacción con el usuario, es claro que al ser una página web con contenido estático, la interactividad es nula como ya se mencionó, por lo anterior, las aplicaciones que se encuentran dentro del marco de la web 1.0 no permiten a los

¹⁷⁶ “Los marcos HTML permiten a los autores presentar documentos con vistas múltiples, que pueden ser ventanas o subventanas independientes. Las vistas múltiples ofrecen a los autores una manera de mantener cierta información visible mientras otras vistas se desplazan o se sustituyen. Por ejemplo, dentro de una misma ventana, un marco podría mostrar un gráfico estático, un segundo marco un menú de navegación, y un tercero el documento principal que puede ser desplazado, o reemplazado al navegar por el segundo marco”.

Información completa en la traducción realizada por POZO JUAN R. Specification, publicada el 24 de diciembre de 1999. La dirección URL de este documento es <http://html.conclase.net/w3c/html401-es/cover.html> La dirección URL del documento original es <http://www.w3.org/TR/1999/REC-html401-19991224> La dirección URL de la última versión de la Recomendación en inglés es <http://www.w3.org/TR/html401>.

¹⁷⁷ Los creadores o dueños de la página web, utilizan los libros de visitas on line o guestbooks con el fin de conocer más sobre sus visitantes y sus opiniones.

usuarios observar cómo funciona, realizar cambios, crear nuevas aplicaciones basadas en las que se encuentran en la página objeto de consulta o realizar contribuciones a la página, “con excepción de las manipulaciones que permite el lenguaje de hipertexto y los hiperenlaces consistentes en poder pasar de una página web a otra”¹⁷⁸.

Gracias a World Wide Web (WWW), en la década de los años noventa, se proporciona facilidad en la consulta de archivos de hipertexto, razón por la cual al hablar de web 1.0, es pertinente e imprescindible, realizar una explicación sobre el concepto de HTML, en sus siglas (Hiper Text Markup Language), también denominado hipertexto, que es el precursor y creador de la Web 1.0. Toda vez que con aquel se empezaron a “escribir los documentos del world wide web”¹⁷⁹.

Ahora bien, una vez explicado lo anterior, es indispensable hacer claridad sobre el concepto de HTML, también denominado formato estático, que es la denominación otorgada a las herramientas contenidas dentro de una página Web, que facilita las transiciones a otras locaciones Web, verbigracia: Como el texto que en la pantalla de un dispositivo electrónico conduce a otro texto relacionado. Existen dos formas de HTML, que permiten el funcionamiento de WWW, estos formatos son: El hipertexto y el lenguaje de marcación. El primero de ellos, es una forma de crear documentos multimedia y vincularlos dentro o entre ellos y el segundo es un método para incluir etiquetas especiales (también conocidas como elementos) que describen la estructura, así como el comportamiento del documento¹⁸⁰.

¹⁷⁸ MÁRQUEZ SOLÍS, S. (2007). *La web semántica y servicios web semánticos*. Universidad Oberta de Cataluña. p. 18.

¹⁷⁹ CAFASSI, E. op. cit. p. 260.

¹⁸⁰ CHELSEA V. Y OTROS. *HTML 4 para dummies*. (1997). Panama: IDG BooksWorldwide. p. 10.

El HTML, es frecuentemente utilizado como lenguaje de marcado para la elaboración de páginas web. Igualmente es definido como aquel lenguaje de programación usado para armar páginas Web¹⁸¹.

En igual sentido, un hipertexto contiene los siguientes elementos: Nodos o secciones, enlaces o hipervínculos y anclajes. Los nodos o secciones, son aquellas partes del hipertexto que contienen información accesible para el usuario. De otro lado, encontramos el concepto de enlaces o hipervínculos, del cual haremos referencia en líneas posteriores, toda vez que su concepto está revestido de vital importancia, sin embargo, dicho sea de paso, los enlaces, como su mismo nombre lo indica son las conexiones, uniones o vínculos que se establecen dentro de los nodos, facilitando la lectura secuencial o no secuencial por los nodos del documento. Y finalmente los anclajes son aquellos puntos de activación de los enlaces¹⁸².

Uno de los conceptos, antes ya mencionado, que resulta fundamental a la hora de hablar de HTML, es el *hipervínculo*, que es el direccionamiento existente en un documento hipertexto que enlaza o conduce hacia otro documento, sea este hipertexto o no. La existencia de los códigos HTML, permiten crear los vínculos o hipervínculos, igualmente, se ha mencionado que “cuando un encabezado Web habla de vínculos, se refiere a transiciones entre locaciones Web, ya sea dentro del mismo documento o con otros documentos en alguna otra parte de la Red”¹⁸³. Otro aspecto importante en relación a HTML, es que a diferencia de los bancos jurídicos

¹⁸¹ SIRI, L. (2000). op.cit. p. 112.

¹⁸² *Hipertexto*. (2015, enero 23). Wikipedia, La enciclopedia libre: <http://es.wikipedia.org/wiki/Hipertexto>

¹⁸³ CHELSEA V. Y OTROS. op. cit.

desarrollados con anterioridad al surgimiento de internet, este carece de limitaciones sobre los elementos incorporados dentro del mismo, toda vez que puede contener datos textuales, dibujos, sonidos e incluso videos relacionados con el tema.

El origen de la palabra hipervínculo reside en la idea de interacción entre los diferentes documentos y el lector, de forma tal que los textos, la voz, las imágenes y los videos ofrecen dicha posibilidad a los usuarios. Paralelamente, debido a la imperante necesidad de procesar altas cantidades de información, se buscó desarrollar la hipermedia y el hipertexto por conducto de un software.

Ahora bien, el formato HTML conocido como el precursor de la web 1.0, en su sentido prístino era un formato totalmente estático, motivo por el cual la página web no se actualizaba, este formato HTML, viene a ser remplazado por la Web 1.5 que se asocia a la tecnología de los años 1997 a 2003, en esta surgen formatos como DHTML, ASP y CSS, que “son páginas web construidas dinámicamente a partir de una o varias bases de datos”¹⁸⁴.

No obstante lo anterior, la web 1.0 y la web 1.5 cuentan con una característica común: la unidireccionalidad de la información, es decir que la generación de información proviene únicamente de su creador, de manera que el usuario no se encuentra facultado para interactuar con los contenidos.

Por tanto, la web 1.5 que “cuenta con contenidos dinámicos, facilitan y proporcionan interacción con el usuario, el sistema presenta respuestas en función de los criterios

¹⁸⁴ TORRES, A. (2009, Mayo 3). *WEB 2.0 Qué es y que nos espera?* Consultada en marzo 5, 2012, WEB 2.0+TIC hacia donde vamos!: <https://aulaticujap.wordpress.com/unidad-i/about/>

introducidos en formularios. Por lo tanto la experiencia con el usuario queda limitada a la aplicación que utiliza»¹⁸⁵.

Pues bien, las páginas en la web 1.5, ofrecen la oportunidad de encontrar contenidos actualizados y variados, en las diferentes visitas que el usuario haga al mismo sitio web.

En lo concerniente a la informática jurídica, cabe anotar que la web 1.0 se ha hecho presente para la creación de varias bases de datos con contenidos legislativos y jurisprudenciales; por ello encontramos páginas web jurídicas como: Norm@civil que es una base de datos jurídica especializada en la recopilación de normativa de naturaleza civil que contiene legislación civil catalana, estatal y jurisprudencia civil del Tribunal Superior de Justicia de Cataluña, entre otros contenidos. Este portal jurídico es elaborado por el Instituto de Derecho Privado Europeo y Comparado de la Universidad de Girona, con el apoyo del Departamento de Justicia de la Generalitat de Cataluña¹⁸⁶.

En igual sentido, en la web 1.0 se encuentra la página del portal jurídico de la empresa Leggio Contenidos y Aplicaciones Informáticas, S.L., este portal de acceso gratuito ofrece en la actualidad, acceso a legislación, jurisprudencia, notas de actualidad, artículos doctrinales, enlaces de interés, guía judicial española, entre otros. Cabe mencionar, que desde el año de 1996, esta compañía ha trabajado en temas de informática jurídica, dando paso a la creación de su página web Noticias

¹⁸⁵ Márquez Solís, S. op. cit. p. 18

¹⁸⁶ INSTITUTE OF EUROPEAN AND COMPARATIVE PRIVATE LAW. *Norma Jurídica*. (2014, Junio 4). <http://civil.udg.es/NORMACIVIL/>

Jurídicas, que con posterioridad se unió a la Editorial Bosch para desarrollar bases de datos jurídicas¹⁸⁷.

Iustel es una base de datos jurídica que ha hecho uso del lenguaje de marcado, cuenta con este portal jurídico, que ofrece además de un compendio legislativo, repertorio jurisprudencial, base de conocimiento jurídico, bibliografía jurídica, revistas generales de derecho y brinda la posibilidad de realizar compras virtuales¹⁸⁸.

En Colombia encontramos bases de datos en formato HTML, una de ellas es la base de datos jurídica de la jurisprudencia de la Corte Constitucional de Colombia¹⁸⁹, en igual sentido, en lo concerniente a las bases de datos de entidades del gobierno, podemos mencionar que existe la página de juriscol, desarrollada e implementada por el Banco de la República, que al igual que el banco de datos de la Corte Constitucional facilita el acceso gratuito a la información.

El Grupo Legis, cuenta con bases de datos, accesibles en su página web, atienden al formato de la web 1.0, toda vez que son de contenido estático, lo que no significa que no ofrecen un contenido jurídico amplio¹⁹⁰.

Otra base de datos jurídica que implementó el formato html, es Portal Jurídico DMS. Esta página web contiene bases de datos jurídicas colombianas y de normativa internacional, sus contenidos serán detallados en la parte final de esta segunda parte

¹⁸⁷ NOTICIAS JURÍDICAS. (n.d.). <http://noticias.juridicas.com/>

¹⁸⁸ PORTALDERECHO. IUSTEL. (2001). <http://www.iustel.com/>

¹⁸⁹ COLOMBIA. CORTE CONSTITUCIONAL. (n.d.). <http://www.corteconstitucional.gov.co/>

¹⁹⁰ LEGIS COLOMBIA. (n.d.). <http://www2.legis.com.co/>

cuando hablemos de las bases de datos que se encuentran en el presente y particularmente en Internet¹⁹¹.

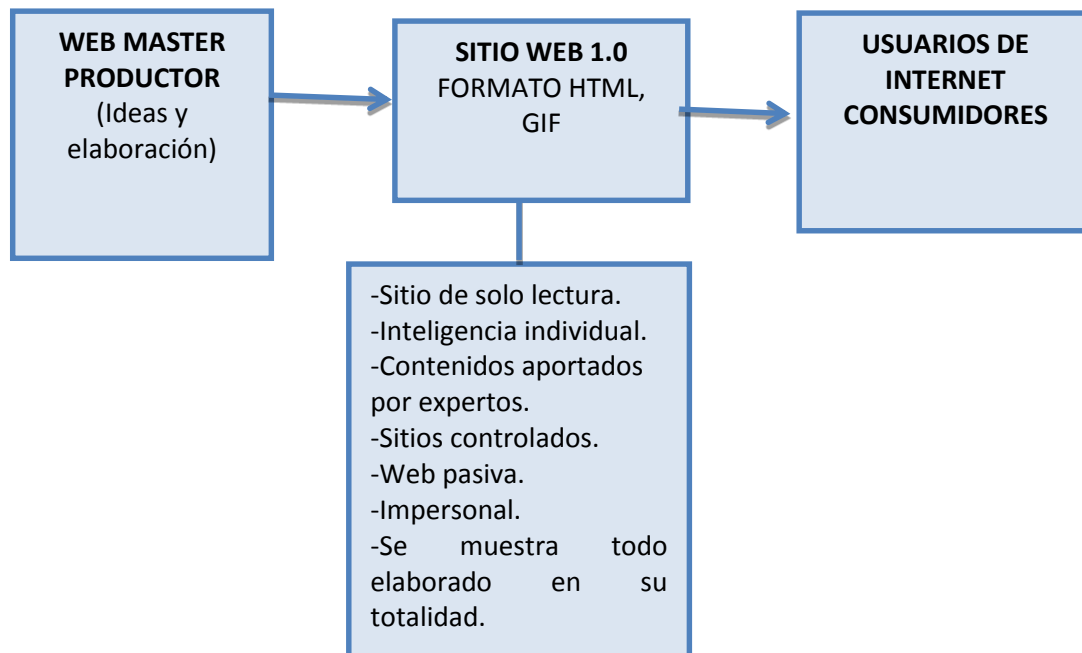
Enfocando la Web 1.0 a la informática jurídica, encontramos que existen páginas en formato HTML, que suministran información jurídica y que en particular ofrecen bases de datos en la gran mayoría de casos con un acceso gratuito. Este conjunto de bases de datos de la web 1.0 serán susceptibles de ser detalladas con exhaustivo detenimiento en los acápites finales de esta segunda parte, toda vez que es imprescindible con anterioridad tener claros los conceptos, funcionalidad, ventajas y desventajas de los buscadores generales que se encuentran igualmente en la web, analizados desde la perspectiva de la informática jurídica documental, con el fin de poder reconocer la importancia de estos bancos de datos jurídicos especializados.

De forma sucinta, podemos sostener que en la web 1.0 la incorporación de audiencias dentro de las bases de datos jurídicas que se encuentran en Internet actualmente, en las cuales además de poder observar todo lo correspondiente con el ordenamiento jurídico y la regulación especial de determinada materia, suministra a los usuarios un contenido considerablemente alto de modelos de audiencias orales, que pueden incluso llegar a contener la totalidad del proceso. Este modelo, es creado dentro de la Red gracias a HTML, por cuanto este ha sido la pieza fundamental que ofrece la posibilidad de interacción de multiplicidad de archivos no solo de video, sino también de texto, voz e imágenes con los usuarios.

Una gráfica que puede llegar a representar la Web 1.0 es la siguiente:

¹⁹¹ Portal Jurídico. (2013, Agosto 1). <http://www.portaldms.com/index.html>

Figura 2. Representación de la web 1.0



Fuente: elaboración propia, con base en conceptos y desarrollo de la web 1.0

2.5. Web 2.0

La Web 2.0, que surge en 2004, a diferencia de la Web 1.0 que cuenta con usuarios pasivos y páginas diseñadas en HTML y GIF, tiene por clave principal el uso de la tecnología AJAX (Lenguajes de programación JavaScript y XML), y una plataforma de colaboración social. Por lo anterior, el sistema antiguo de información estática, ha sido remplazado por una Web dinámica que otorga la posibilidad a los usuarios de publicar sus propios contenidos. De manera que, en la Web social, al no existir absoluta confidencialidad en la información, resulta fácil acceder a la misma por

parte de los usuarios al momento de emprender una búsqueda sobre un tema determinado.

Para corroborar lo anteriormente expuesto, basta con observar el surgimiento de la Web 1.0 en la cual existen páginas Web con características estáticas desde el punto de vista del usuario, pero con libertad para el creador o administrador con relación a su contenido libre. Aún más claro, es el contenido insubordinado en la denominada Web 2.0 también conocida como Web social, que permite la participación e interacción entre los usuarios en gran variedad de sitios. La Web 2.0 es definida como aquella que permite dinamizar la información gracias a la interacción y el valor que le puede otorgar cada usuario que hace uso de la misma y que participa de forma interactiva y transparente en el intercambio de conocimientos, igualmente es entendida como “la filosofía de maximización mutua de la inteligencia colectiva”¹⁹².

Dentro de la Web 2.0 y sus formatos, podemos encontrar: Mensajería instantánea, Blogs y Vblogs que no son otra cosa que video blogs, foros de internet, chat, compartimiento de ficheros tales como música; fotografía; videos; juegos, podcast (blogs de audio), wikis que son sistemas de generación de contenido en línea creados y editados por los usuarios conectados al mismo sitio web y Redes Sociales, entre otros.

Como características de la web 2.0, podemos mencionar que su contenido es colaborativo, por lo cual la información es compartida a los usuarios que socializan dentro de los diferentes servicios que en ella se ofrecen, hecho que contribuye al

¹⁹² JOYANES AGUILAR, L. (2009). *Redes sociales: la nueva revolución de la era de facebook*. España: Joxman Editores Multimedia. p. 59.

incremento de la inteligencia colectiva, por lo cual otra de las características con las que cuenta la web 2.0 consiste en el intercambio y reutilización de la información en ella contenida.

Esta web, que cuenta con un contenido colaborativo, “podemos decir que es el modelo que se está imponiendo poco a poco, y que hace uso de las capacidades adquiridas en la generación anterior, permite que la experiencia del usuario con la Web mejore espectacularmente. En esta generación, las aplicaciones van más allá de la mera interacción entre aplicación-usuario-aplicación, ahora ellas son el mecanismo que permiten que se produzca interacción entre usuario-usuario llegando a crearse un entorno de contenido colaborativo, en donde, el usuario es un participante más en la creación del contenido que aparece en la web. Ideas como los wikis, los blogs, etc. han cambiado la forma en la que el usuario interactúa con la web haciendo de esta un lugar mucho más rico e interesante para trabajar”¹⁹³.

Una característica determinante, a la hora de hablar de la web 2.0, es el uso del formato XML (eXtensible Markup Language), que es un lenguaje de marcas más potente que el HTML, que realiza un aporte semántico a la información y ofrece independencia, integración y en especial el uso de cualquier medio que cuente con acceso a Internet.

Dentro de las ventajas relevantes del formato XML, encontramos que cuenta con estándares que permiten la integración e intercambio de datos indistintamente del lugar donde se hallen, en igual sentido el lenguaje XML, “a) es un estándar

¹⁹³ MÁRQUEZ SOLÍS, S. (2007). *La web semántica y servicios web semánticos*. Universidad Oberta de Cataluña. p. 8.

internacional; *b*) es una tecnología abierta, no propietaria, independiente de plataforma y sistema operativo; *c*) se centra en la separación de contenido frente a la presentación; *d*) tiene gran potencia para construir vocabularios aplicables a cualquier tipo de documento; *e*) se orienta al tratamiento, transmisión o intercambio de todo tipo de documentos o datos; *f*) permite reutilizar textos y datos ya existentes en unos documentos para la confección de otros nuevos; *g*) aporta potentes mecanismos para la búsqueda y recuperación de la información; *h*) permite cualquier dispositivo de consulta: navegador, PDA, teléfonos móviles, etcétera; *i*) es flexible para permitir la integración con otros sistemas y con objetos no XML; *j*) es capaz de dotar de seguridad y fiabilidad al documento.

XML permite generalizar la estructura de un documento asignando al mismo tipo de documentos la misma DTD (Document Type Definition). Por su parte, los esquemas XML ofrecen un mecanismo más potente y expresivo y permiten definir con mayor riqueza y complejidad estructuras internas, así como hacer uso de otros estándares XML acompañantes, como Xpath, XSL, XML Namespaces, Xlink, etcétera. Martínez González et ál. (2001) evidencian el interés intrínseco de estas tecnologías para ofrecer una solución a las complejas relaciones entre documentos jurídicos señalando cómo los estándares asociados a XML, Xlink, Xpointer, Xpath proporcionan las herramientas necesarias para direccionar los fragmentos de los documentos y modelar enlaces tan complejos como se desee. La asociación de Xlink y Xpointer permite asociar semántica a los enlaces, crear enlaces multidireccionales, asociar múltiples recursos en un mismo enlace y crear bases de enlaces independientes de los documentos¹⁹⁴.

¹⁹⁴ ALVITE DÍEZ, M. L. (2009). op.cit. p. 37.

Otra característica de la web 2.0, consiste en la facilidad que proporciona para el acceso a la información, bien sea para compartirla con otros usuarios de la web o para estudiarla en sus quehaceres jurídicos, investigaciones o sencillamente para informarse sobre asuntos de novedad del derecho.

Este fenómeno de la Web Social, no ha sido ajeno a nadie que cuente con acceso a Internet, es tan trascendente su surgimiento como los cambios que este conlleva, son estos, la socialización de la Web y el cambio de la información estática a la utilización de medios de comunicación en masas. A todos nos resulta familiar hablar de Twitter.com, Facebook.com, Youtube.com, Wikipedia.org, Amazon.com. En adición a lo anterior, existe un fenómeno que ha difundido ostensibles cantidades de información: Los denominados *blogs*.

Los *blogs*, son páginas Web personales, libres, públicas e interactivas, organizadas de forma cronológica, dentro de los cuales el creador o Blogger se expresa sobre cualquier tema de interés, gracias a la utilización de una herramienta de edición, en ellos existe la posibilidad por parte del usuario de realizar comentarios, enlazarla e incluso escribirle al autor del blog. Es común, encontrar dentro de los mismos, publicación de noticias novedosas, de igual forma, son utilizados como elementos de publicación, de conversación o de generación de influencia.

Como consecuencia para la informática jurídica, en la Web 2.0 existen, videos, foros de internet e incluso blogs jurídicos, no obstante que el uso de estos últimos, en la

actualidad continúa siendo restringido en las facultades de derecho, y su utilización como herramienta para los usuarios en general, aún está en proceso de desarrollo.

Cabe señalar, que los blogs jurídicos, con el fin de lograr reconocimiento y resultar una herramienta útil, además de generar un espacio para la discusión entre los usuarios de internet, en la actualidad suministran y analizan información generada en la actividad profesional, igualmente opinan y analizan casos en concreto, reproducen, sistematizan o complementan información jurídica.

En el mejor de los casos, existen blogs unidireccionales y poco colaborativos, pero que se dirigen a otro tipo de iniciativas como los blogs de bibliotecas de derecho, aquí “podemos encontrar el blog de la Biblioteca de la Facultad de Derecho de la Universidad de Barcelona, el del campus de Almagro de la Universidad Castilla-La Mancha o el de la Universidad Complutense de Madrid. Su ánimo suele ser informativo sobre actividades domésticas (administrativas), pero también pueden recopilar información de interés jurídico sobre acontecimientos externos, como reproducción de noticias relacionadas con el mundo del derecho. Estos blogs suelen contener numerosos enlaces a otros blogs de abogados, portales de empresas editoriales y todo tipo de recursos jurídicos. El blog de la Universidad Complutense toma una inclinación fundamentalmente difusora al ofrecer una relación de enlaces a los artículos publicados por sus profesores”¹⁹⁵.

A pesar de lo anterior, la divulgación de información jurídica en los blogs es limitada, toda vez que el derecho continúa con su desarrollo habitual en las

¹⁹⁵ RAMÍREZ CLEVES, G. (2010). *Los blogs jurídicos y la web 2.0 para la difusión y la enseñanza del derecho*. Bogotá: Universidad Extenado de Colombia. p. 74.

universidades, los congresos, conferencias y despachos judiciales. “Apenas tienen repercusión los blogs jurídicos cuando de lo que se trata de comentar es un suceso relacionado con el uso de la tecnología o el devenir del derecho, y tampoco existe una comunidad de juristas significativa que se haya incorporado al espacio de Internet”¹⁹⁶. Incluso, en la actualidad son pocos los dedicados a la recopilación del ordenamiento jurídico, por lo cual la existencia de documentación jurídica se encuentra centralizado en contenidos web que en la mayoría de los casos son de acceso restringido.

Dentro de la Web 1.0 y 2.0, con el fin de encontrar información útil, los usuarios emplean los buscadores como herramientas de selección de información, dado que arrojan las más de las veces un alto contenido de información dentro del resultado de la búsqueda, sobre la cual el usuario debe desarrollar la actividad de discriminación y verificación de la confiabilidad de la misma, pudiendo caer dentro del curso de este trabajo, en errores que pueden llegar a alterar los resultados reales, oportunos, pertinentes y veraces esperados¹⁹⁷.

Para tales efectos a continuación se estudiará el fenómeno de los buscadores o *search engine*.

No obstante lo anterior, teniendo en cuenta que la capacidad de almacenamiento se incrementó de forma ostensible, pese a que desapareció el problema que se presentaba en el pasado, el surgimiento de la World Wide Web, la creación de la web 1.0, la web 1.5 y web 2.0, trajo consigo la inflación de información, también

¹⁹⁶ Ibídem. p. 101.

¹⁹⁷ SIRI, L. (2000). op.cit. p. 19.

denominada “hiperinflación” anteriormente ya mencionada, que consiste en el alto contenido de información documental con que contamos en Internet, de esta forma, pese a que el usuario tenga los conocimientos técnicos necesarios, puede ser víctima de la desinformación a causa de este fenómeno, incluso en los bancos de datos especializados.

El problema real, consiste en que es demasiado difícil encontrar lo que se necesita, a diferencia de otros aspectos de Internet, este problema afecta a todos, desde novatos a expertos o desde usuarios a desarrolladores, en especial porque en la era que vivimos actualmente cualquier persona puede en principio publicar de forma libre en la Web.

Ahora bien, para los efectos de la informática jurídica, a la luz de la confiabilidad de la información, resulta necesario tener presente que la información contenida en la web 2.0 es susceptible de ser revisada, toda vez que como se explicó en párrafos anteriores, cualquier persona sea o no experto en algún tema puede realizar aportes a la web 2.0 que en gran cantidad de ocasiones no se encuentran respaldados o son erróneos.

De acuerdo con lo anterior, las aplicaciones que ofrece la web 2.0, a diferencia de la web 1.0, son códigos abiertos de manera que todos los usuarios cuentan con la posibilidad de crear nuevas aplicaciones, modificar las ya existentes, realizar adiciones. En igual sentido, en la web 2.0 son todos los usuarios, quienes comparten la información, motivo por el cual ellos son los creadores de la denominada “inteligencia colectiva”.

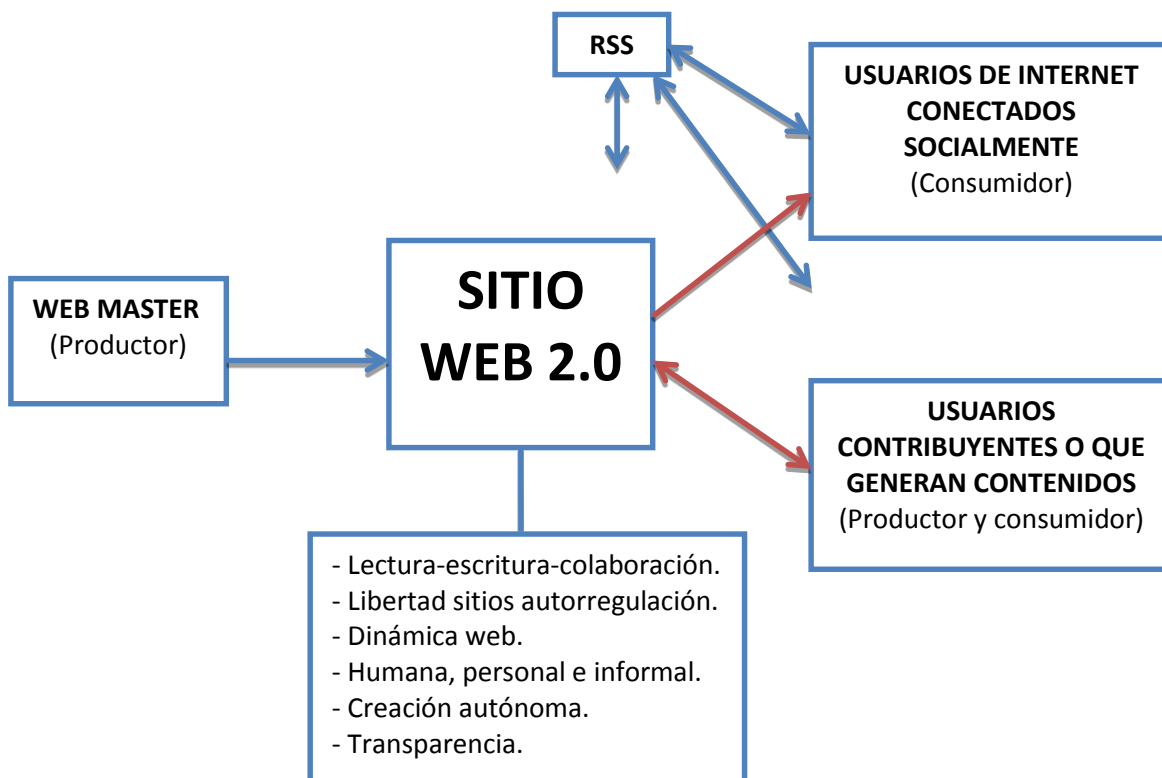
Una característica significativa, reside en la posibilidad con la que cuenta cualquier usuario web 2.0, para publicar libremente contenidos, razón por la cual, a diferencia de la web 1.0, estos usuarios son además de consumidores en gran cantidad de casos, también productores de información. En este punto, cabe mencionar, que a pesar de la posibilidad de ejercer su libertad de expresión y de no existir barreras para transmitir conocimiento, esta facultad, trae consigo, entre otros, problemas de veracidad en los aspectos intrínsecos, desde el punto de vista de la informática jurídica y del documento jurídico.

De otra parte, en términos de interacción, no para hacer referencia de la web 2.0 en general, sino tratándose de aquellos bancos de datos jurídicos que se encuentran en la web y son elaborados de forma confiable con formatos y aplicaciones web 2.0, se reconoce la mayor utilidad, toda vez que al ofrecer un manejo dinámico de la información, actualizada de forma constante, la interacción con el usuario se facilita logrando una retroalimentación, en vista de la posibilidad con la cual ellos cuentan para comunicar sus impresiones y difundir la información.

Es de anotar, que no siempre que hablamos de la web 2.0, hacemos referencia a las redes sociales, toda vez que como acabamos de ver, existen plataformas 2.0 dentro de las cuales los usuarios juegan un papel importante y son tenidos en cuenta para su presentación y mejoramiento, tal es el caso de bases de datos jurídicas presentadas en Internet en la web 2.0, que hacen uso de la página para interactuar con aquellas personas que los consultan, de forma que tendrán la posibilidad de dejar comentarios, interactuar con el web master y referenciar sus contenidos.

A continuación, podemos observar la gráfica representativa de la web 2.0:

Figura 3. Representación de la web 2.0



Fuente: elaboración propia con base en conceptos y desarrollo de la web 1.0

2.6. Antecedentes de los Search Engine generales

El surgimiento de la búsqueda en los medios de información o electrónica, se remonta al año de 1945, cuando Vannevar Buch (maestro de la computación analógica) publicó “As We May Think” «Como podríamos pensar» en la revista

Atlantic Monthly.¹⁹⁸ En este artículo, surgió la idea sobre el desarrollo de las nuevas tecnologías, años después materializada por Ted Nelson Y Douglas Engelbart, quienes gracias a las reflexiones posteriores a este artículo, lograron desarrollar la tecnología del hipertexto.

Con la publicación de su artículo, Vannevar Buch,¹⁹⁹ es considerado un visionario, toda vez que predijo varios de los avances tecnológicos con los que contamos en la actualidad, describió la llegada de varios dispositivos tales como: La máquina pensante, que en su momento sirvió para realizar los cálculos matemáticos que hoy en día son efectuados por las calculadoras. El Vocoder, que es un lector de voz que

¹⁹⁸ The Atlantic Monthly o Atlantic, es una revista que tiene origen en el año de 1857 en Boston, cuyo contenido principalmente es literario y cultural.

En ella existen artículos sobre temas políticos al igual que de asuntos internacionales y críticas literarias. Dentro de esta prestigiosa revista, se han publicado artículos de autores de conocida trayectoria e incluso textos de Martin Luther King, Jr. y textos teóricos que han propiciado la creación de nuevas tecnologías como la del hipertexto. Su sede permaneció en Boston hasta que en el año 2005, una vez adquirida la revista en 1999 por el dueño del Grupo Nacional de Periódicos con sede en Washington D.C., se tomó la decisión de trasladar a esta última ciudad la revista con el objetivo de unirse a las divisiones de publicidad y circulación de la nueva compañía adquirente y disminuir los costos que acarreaba tener la revista en Boston.

¹⁹⁹ Alumno aventajado del Tufts College su tesis en 1913 mostraba su Profile Tracer, invento que servía para medir distancias entre terrenos desiguales. En 1919, se une al Departamento de Ingeniería Eléctrica del MIT, donde permaneció 25 años, 12 de ellos como Decano. Entre sus proyectos, dispositivos ópticos y de composición fotográfica, almacenamiento y recuperación de microfilms. Con su compañero de habitación en Tufts Laurence K. Marshall y el científico Charles G. Smith, fundan en Cambridge, Massachusetts en 1922 una compañía con el nombre de American Appliance Company. Posteriormente convertida en Raytheon, principal contratista en materia de Defensa del Gobierno de los EE.UU. y muy ocupado tras los ataques del 11 de septiembre. (Entre sus productos: detectores de agentes químicos, visores infrarrojos, ciberseguridad, traductores árabe-inglés...).

En 1939 es nombrado presidente del Carnegie Institute de Washington; Director del National Advisory Committee for Aeronautics; y nombrado por el presidente de EE.UU. Director de la Office of Scientific Research and Development encomendándole la gestión de los más de 6000 científicos envueltos en proyectos relacionados con la defensa de EE.UU. en la II Guerra Mundial, entre ellos, la creación de la Bomba Atómica.

Para la realización de cálculos repetitivos diseña el Analizador Diferencial, primer computador analógico condenado a la obsolescencia por la aparición de la computadora digital. Este dispositivo permitía la resolución de ecuaciones de hasta 18 variables y sirvió para cálculos balísticos entre otros.

En 1947 es nombrado director de AT&T y en 1948 director de Merck and Co.

traslada lo dictado en forma oral a formato escrito. Hoy en día este dispositivo es el software denominado Voice Type.

Para el tema que nos concierne, encontramos la propuesta de creación del dispositivo Memex, que a pesar de no haberse desarrollado, influyó de forma considerable en lo que hoy en día conocemos de Internet e hipertexto. Dentro de la idea de Memex, se planteó la posibilidad que el usuario contara con el almacenamiento de la totalidad de sus archivos: Libros, grabaciones y comunicaciones, de forma que pudiese emprender su consulta de manera rápida y flexible. La idea física de lo que sería Memex, consistiría en una superficie en la cual se ubicarían dos monitores, manejados mediante un lápiz óptico y un escáner, que en su parte interna almacenaría la información indexada compuesta tanto por texto como por imágenes, siguiendo estándares universales.

Con la idea de Memex, se planteaba la posibilidad de que cualquier persona que hiciera uso de su tecnología pudiese consultar información y crear rutas de almacenamiento, con la posibilidad de otorgarles un nombre de acceso propio de forma que se pudiese consultar y utilizar de nuevo por parte de otros usuarios. Para la tarea de instalación de rutas de almacenamiento se planteó la formación de los llamados “Trail Blazers” usuarios cuya misión consistiría en crear el acceso propio para la información consultada con el fin de dejar abierta la eventualidad de una posterior consulta.

En este punto, resulta necesario establecer que la propuesta de la creación del dispositivo Memex, no tuvo un desarrollo entendido tal como se había planteado en

la idea de Vannevar Buch, no obstante, este concepto prístino fue la base de la creación del hipertexto e Internet, por parte de Douglas Engelbart, Ted Nelson y Tim Berners Lee.

Dicho lo anterior, es importante mencionar que los trabajos desarrollados por Douglas Engelbart, Ted Nelson y Tim Berners Lee, fueron de vital importancia no solamente para lo relacionado con la creación del hipertexto, sino también para la creación de las redes de computadores y el arribo posterior de Internet. Por ello explicaremos, a continuación, cuáles fueron los aportes de cada uno de ellos.

En cuanto a los aportes de Douglas Carl Engelbart, quien es el inventor estadounidense, conocido por haber co-inventado el mouse del computador,²⁰⁰ podemos afirmar que es pionero de la interacción humana con los computadores, dentro de lo cual se incluye el hipertexto y los computadores en red. Una vez terminada la Segunda Guerra Mundial, Englebart renunció a su trabajo como ingeniero e inspirado por el artículo de Vannevar Bush, trabajó en la versión previa de Internet, es decir, la red militar ARPANET, en otros acápites ya estudiada. Engelbart realizó importantes aportes en el diseño del primer sistema en línea, denominado On-Line System (NLS) en Stanford Research Institute. De hecho, en su oficina se encontraba el segundo nodo de la red.

Gracias a su arduo trabajo y en especial al gran aporte tecnológico que permitió la interacción de las personas con el computador, al crear la interfaz que realizaba la

²⁰⁰ En 1968, en una conferencia de expertos en informática, Doug Engelbart hizo la presentación donde realizó la primera demostración pública del mouse, en la cual incluyó una conexión en pantalla con su centro de investigación, esta fue además el primer vídeo-conferencia de la historia.

unión de los dos, en 1997, le otorgaron el Premio Lemelson-MIT. De igual forma, fue titular de varios premios tales como el Memorial Yuri Rubinsky, el Premio Turing, la Medalla John von Neumann y la Medalla British Computer Society's Lovelace²⁰¹.

En lo concerniente a los aportes que realizó Theodor Holm Nelson o Ted Nelson, es relevante mencionar en primera instancia que fue quien acuñó el término “hipertexto”. Al igual que Engelbart, de forma independiente también inventó el hipertexto de los computadores y pudo entender como los computadores se transformarían en la página impresa. Todo lo anterior, gracias a la fundación de Xanadu, que es el primer equipo del proyecto de hipertexto en 1969. Ya en 1960, él había iniciado el diseño de los documentos electrónicos, con el fin de realizar una distribución digital de los mismos y con la confianza que su idea sería prontamente trasladada a la pantalla digital. Ted había manifestado el desconocimiento del proyecto de Engelbart, del cual tuvo conocimiento en el año de 1965, fecha en que publicó su primer diseño.

Finalmente encontramos a Tim Berners-Lee, quien es conocido por ser el fundador y director de World Wide Web Consortium, Tim escribe el código de la web que conocemos en la actualidad e incluye relaciones entre los documentos y enlaces escritos, de igual forma, estudia las obras de Ted Nelson, esto aunado al hipertexto, le dio paso a lo que hoy en día conocemos como buscador o search engine.

²⁰¹ DOUG ENGELBART INSTITUTE. (2015, enero 25). *Curriculum Vitae*. Consultada en enero 25, 2015, <http://www.doungengelbart.org/about/cv.html>

El padre de la Web, Berners-Lee, quien desarrolló ideas importantes que hoy forman parte de la base de la Web, creó junto con su grupo de trabajo el lenguaje HTML o (Hiper Text Markup Language) explicado anteriormente y de forma adicional, realizó un importante aporte creando el sistema de localización de objetos en la web URL (Uniform Resource Locator).

Con las anteriores herramientas, otorgadas por Berners-Lee y su grupo de trabajo, se da un paso importante para que quien se cree que es el padre de la búsqueda moderna, Gerard Salton, pudiese desarrollar en 1968 un sistema para el análisis mecánico de recuperación de texto denominado SMART (System for Manipulation and Retrieval Text), desde esta fecha, los especialistas en sistemas han trabajado de la mano de bibliotecarios y documentalistas, con el fin de desarrollar un software capaz de recuperar información de forma rápida, con el uso del modelo de recuperación de información de espacio vectorial.

Antes de la creación y el surgimiento de la World Wide Web, encontramos que existieron buscadores que tenían como fin organizar la cantidad de información contenida en Internet. El precursor de esta idea fue Archie de la Universidad McGill en 1990, quien posteriormente fue precedido por WAIS y Gopher. Pese a que estos tres sistemas, fueron anteriores al surgimiento de la WWW, continuaron indexando la Web y el resto de Internet durante varios años después del surgimiento de la WWW.

Posterior al crecimiento de la Web, surgieron los buscadores y los directorios web, con la finalidad de localizar con mayor facilidad la información. El primer buscador

Web completamente de texto fue WebCrawler en el año de 1994. Al decir que WebCrawler era un buscador completamente de texto, hacemos referencia a que los anteriores, ofrecían a los usuarios únicamente los títulos de las páginas Web. Al igual que WebCrawler, es importante mencionar que el primer buscador que contó con éxito comercial fue Lycos, creado en 1993, no obstante en el año de 1995, arriban a Internet Yahoo y Altavista, quienes se posicionan en el mercado como los líderes de la industria. Que en 1998, fueron desplazados por GOOGLE.

Como ya mencionamos, los buscadores en un principio se enfatizaron en su capacidad de almacenamiento, circunstancia que se vio sustituida sustancialmente durante los primeros años del 2000, motivo por el cual, la problemática cambió y suscitó un nuevo planteamiento que consistía en organizar esa cantidad de información según su relevancia.

Entre los varios modelos de recuperación de información existen dos en particular que cobran importancia en este punto histórico: El modelo booleano y el modelo vectorial. En el modelo booleano, la recuperación consiste en la obtención de todos aquellos documentos que cumplan a la luz de la lógica con lo expresado en la consulta, para tales efectos, los documentos son descritos mediante un conjunto de términos de indexación y se utilizan los operadores propios del álgebra de Boole²⁰² (AND, OR Y NOT).²⁰³

²⁰² “Un álgebra de Boole es una entidad matemática formada por un conjunto que contiene dos elementos, unas operaciones básicas sobre estos elementos y una lista de axiomas que definen las propiedades que cumplen las operaciones.

Los elementos de un álgebra de Boole pueden llamarse falso y cierto o, más usualmente, 0 y 1.

De este modo, una variable booleana o variable lógica puede tomar los valores 0 y 1.

George Boole, su creador, formalizó matemáticamente la lógica en 1854, cuando definió lo que hoy se conoce como álgebra de Boole.

Las operaciones booleanas básicas son las siguientes:

El modelo booleano, presenta ventajas toda vez que emprende búsquedas con la restricción de descriptores, discrimina la información y le ofrece al usuario de forma depurada los resultados de la búsqueda, no obstante, no calcula el grado de relevancia para los documentos que se arrojan como resultado y eventualmente puede ocasionar búsquedas que entregan pocos resultados debido a la inclusión de demasiadas operaciones AND o por el contrario arrojar gran cantidad de documentos debido a la mala utilización de la negación o el uso excesivo de OR.

En búsqueda de otro modelo de recuperación de la información, Gerard Salton presentó un modelo de espacio vectorial con el fin de realizar la indización y clasificación de los documentos y su posterior recuperación, cada uno de los documentos es representado como un vector y a su vez cada componente del vector representa un término denominado descriptor.

El modelo vectorial da solución a algunos de los problemas, sin embargo dicho sea de paso, hay otros inconvenientes que continúan subsistiendo. Los aspectos que viene a solucionar el modelo vectorial, son el filtrado, la recuperación, el indexado y

-
- La negación o complementación o NOT, que corresponde a la partícula no y se respresenta con una comilla simple ('). Así, la expresión x' denota la negación de la variable x y se lee "no x "
 - El producto lógico o AND, que corresponde a la conjunción y de la lógica y se representa con el símbolo \cdot . Así, si x e y son variables lógicas, la expresión $x \cdot y$ denota su producto lógico y se lee " x e y "
 - La suma lógica u OR, que corresponde a la conjunción o y se representa con el símbolo $+$. Así, la expresión $x + y$ denota la suma lógica de las variables x e y , y se lee " x o y ".

PEIRON GUARDIA, M., & SÁNCHEZ CARRACEDO, F. (2002). *Los circuitos lógicos combinacionales*. Barcelona: UOC. p. 10.

²⁰³ Verbigracia: Un documento es relevante sí y solo sí la búsqueda contiene:

- Una palabra: contiene la palabra.
- A and B: contiene las palabras A y B.
- A or B: Contiene A ó B.
- A AND NOT B: Contiene A y no B.

en especial el cálculo de relevancia de información que no se encontraba en el modelo booleano.

El modelo vectorial, representa a los documentos con el uso de un lenguaje natural y utiliza vectores con el fin de obtener un resultado encaminado a señalar la relevancia del documento dentro de la búsqueda concreta, que se obtiene producto de la diferencia de los ángulos para así arrojar un valor que puede determinar la coincidencia del documento con las palabras clave.

El sistema vectorial se obtiene por conducto de la aplicación de la conocida fórmula matemática para hallar el coseno del ángulo entre dos vectores.²⁰⁴ Cabe señalar que al igual que el modelo booleano, presenta algunos inconvenientes relacionados principalmente con tres limitaciones a saber:

La primera consistente en que en ocasiones se puede arrojar dentro del resultado de la búsqueda documentos que no tienen que ver con el objeto de la misma, pero que si guardan coincidencia con alguna parte o palabra contenida en el documento.

En segundo lugar, se puede presentar la polisemia, cuando la palabra objeto de búsqueda es susceptible de estar contenida en dos o más nociones con diferente acepción.

$$\cos \theta = \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{\|\mathbf{v}_1\| \|\mathbf{v}_2\|}$$

²⁰⁴

donde:

- teta es el ángulo entre v1 y v2
- v1 es el primer vector
- v2 es el segundo vector
- . representa el producto
- $\|x\|$ representa la magnitud del vector x

Finalmente, encontramos que los documentos con mayor contenido, quedan representados en menor medida, toda vez que cuentan con menos valores en común y una dimensión más grande.

Es importante señalar, que el modelo vectorial fue utilizado por primera vez en el sistema SMART, de recuperación de información mencionado con anterioridad. Desde los años 60's, gracias al desarrollo realizado por Gerard Salton del sistema SMART (System for the Mechanical Analysis and Retrieval of Text), los ingenieros trabajaron en conjunto con bibliotecólogos y especialistas en documentación con el fin de desarrollar un software eficaz en bases de datos que contuviera miles de libros y artículos. Estos esfuerzos concluyeron en dos resultados concurrentes.

Dentro de este contexto, el primer resultado, que se realizó con la implementación del sistema booleano, consistió en la indexación, la descripción de documentos gracias a la utilización de palabras claves y la posterior organización de estos dentro de una base datos sistemática denominada *thesaurus*, ya explicada en la primera parte de esta tesis que contempla el tratamiento de la informática jurídica en el pasado.

Es menester aclarar que en ninguna instancia se puede considerar el thesaurus como un diccionario de sinónimos y antónimos y que este atiende a lo definido con anterioridad. “Con el empleo de estos programas, el usuario, quien será usualmente

un experto, podrá llevar a cabo una investigación compleja usando los operadores booleanos (AND, OR, NOT, y así sucesivamente)”²⁰⁵.

El segundo resultado que surgió a partir de SMART, fue la creación de programas como Lexis Nexis, obtenidos a partir de la creación de procesos de automatización, con la utilización de un computador que realizaba la tarea de comparar las palabras objeto de la búsqueda con los documentos contenidos en las bases de datos. En programas como el anteriormente mencionado, el computador arrojaba los resultados de la búsqueda en atención al criterio de relevancia, de manera que se otorgaba la posibilidad de emprender búsquedas más avanzadas que podían limitarse por razones de tiempo o proximidad de la palabra clave.

Una de las grandes ventajas, que representa este último resultado, consiste en que gracias a la automatización, se prescinde de la engorrosa consulta de bases de datos manuales, de forma que cualquier persona puede acceder a ellos con la inclusión de las palabras clave en las bases de datos digitalizadas que se encuentren disponibles para emprender una búsqueda.

Pese a lo anterior, este modelo automatizado de búsqueda cuenta con algunas desventajas, tal es el caso de los bancos de datos que contienen información de diferentes áreas o temas, verbigracia: bases de datos que recogen información indiscriminada sobre temas de actualidad, ciencia, derecho, compras, medicina, etc.

En ellas se van a presenciar problemas de léxico, que dicho sea de paso, incluso en

²⁰⁵ “Some automated the work of librarians, who index documents within a subject area, describe those documents with keywords, and then compile those keywords in a database called a *thesaurus* (not to be confused these programs, the user (usually an expert) can perform complex research using Boolean operators (*and, or, not, and so on*)”. GIRARD, B. (2009). *The Google Way: How One Company is Revolutionizing Management as We Know it*. San Francisco: No Starch Press, Inc. p. 13.

los buscadores de la actualidad aún persisten, toda vez que al no ser estas bases de datos especializadas sobre un tema, incurren en polisemias, sinonimias, creando mayor volumen de información y ruido informático.

Los problemas planteados, se pueden aminorar de forma considerable con la creación de bases de datos especializadas, que al concretizar el tema objeto de organización, si se introducen palabras clave, los resultados de la búsqueda serán más limitados y pertinentes, de forma que el usuario ya tendrá mucho más discriminada la información y no cuestionará la procedencia de la misma. Este tipo de bases de datos se crearon en el pasado, tal como ya lo mencionamos dentro del primer acápite de esta tesis, sin embargo, son pocas las bases de datos de esta naturaleza que reposan en Internet y que además proporcionan facilidades en la inclusión de palabra clave, para obtener la información pretendida, dado que su búsqueda se encuentra restringida y este tipo de programas de búsqueda especializada en la web, no fue el centro de atención, ya que lo que se pretendió fue organizar la totalidad de información contenida en la World Wide Web.

Con el anterior panorama, debido a la cantidad de información contenida en la Web, bien fuera 1.0 o 2.0, surgió la necesidad de crear *motores de búsqueda*, esto con el fin de contribuir a mejorar la navegación del usuario y de igual forma lograr la organización y eficacia de los contenidos de Internet. A partir del año 1993 surge el primer search engine denominado “Wandex”; desde entonces, aparecen en la web multiplicidad de buscadores, que, en la totalidad de los casos, arrojan información contenida en Internet de la Web 1.0 y 2.0.

En el afán por organizar el contenido de Internet, esta tarea se tornó cada vez más difícil, puesto que la información contenida en la web, aumentó de forma vertiginosa y trajo consigo gran cantidad de información, por lo cual su organización y discriminación se tornó más difícil, en consecuencia, la calidad de los resultados de la búsqueda se deterioró.

A causa del deterioro en los resultados de la búsqueda, algunos buscadores desaparecieron, por su parte otros buscaron solucionar este inconveniente, para tales efectos limitaron los resultados de la búsqueda y redujeron el tamaño de sus bases de datos, otros buscadores como yahoo, adoptaron la utilización de los thesaurus, de forma que lograron realizar una clasificación y categorización de los sitios posicionándolos según el tema, gracias a la solicitud que realizaba el web master, que consistía en la especificación de la categoría de su sitio web por medio de la utilización de palabras clave, el motor de búsqueda de yahoo controlaba la relevancia de la misma y así se definía su ubicación.

2.7. Concepto de Buscador

En términos generales, podemos afirmar que un buscador es una herramienta informática que tiene por finalidad ayudar a emprender búsquedas sobre temas particulares en atención a diferentes criterios, en otras palabras los buscadores son sistemas automatizados que buscan recuperar la máxima cantidad de información contenida en la web o en una base de datos específica.

Ahora bien, en este punto corresponde introducir el concepto de metabuscadores, que para efectos de los buscadores generales cobra una gran importancia toda vez que estos cumplen la función de unificación de la información relevante dentro de los resultados de búsqueda de cada buscador general, motivo por el cual carece de una base de datos propia, bien sea indexada como Google o temática como Yahoo!.

En otras palabras para aludir al concepto más genérico de un buscador, podemos afirmar que un metabuscador es el buscador que incorpora un conjunto de buscadores. Algunos ejemplos de metabuscadores son: Dogpile, Aleyares, Metacrawler, entre otros. Estos metabuscadores presentan ventajas, como ampliar el espacio de búsqueda y en algunos casos mostrar la posición de la web.

Dentro de esta acepción amplia, cabe mencionar que existen varias clases de buscadores, de forma enunciativa son tres a saber: Los buscadores automáticos o generales, los buscadores especializados y los buscadores temáticos. Estudiaremos en lo sucesivo los buscadores generales y una vez analizado en su totalidad, continuaremos más adelante con las definiciones y discusiones en torno a los buscadores especializados y los buscadores temáticos.

Mencionado lo anterior, podemos definir a los buscadores generales, como aquellas bases de datos con contenidos indeterminados que tienen por finalidad principal ofrecer a los usuarios la máxima cantidad de información contenida en la web y facilitar la búsqueda de enlaces web dentro de WWW de forma automática y expedita. Por lo cual al introducir las palabras claves, arrojan toda la información

relacionada con el objeto de la búsqueda sin distinción alguna sobre la especialidad o el tema concreto que se pretende encontrar.

Otra definición de buscadores generales, sostiene que

Son herramientas de búsqueda que recopilan todo tipo de información y que se caracterizan por contar con una inmensa base de datos, un gran ancho de banda para responder a millones de consultas simultáneas realizadas por los usuarios y una interfaz muy simple. Se tratan de programas que recorren la WWW con una periodicidad determinada, recabando información sobre las páginas web alojadas en los distintos servidores. La base de datos se auto construye con los datos obtenidos en estas visitas, sin que medie ningún tipo de intervención humana que haga un análisis o valoración de las páginas disponibles.

El orden en que se obtienen los resultados al efectuar una búsqueda en uno de estos programas buscadores depende de cada uno de ellos, puesto que utilizan distintos algoritmos que establecen criterios mediante los cuales se realiza la ordenación de los resultados. (...)

Sobre estos recursos es importante tener en cuenta que, en contra de la costumbre generalizada en los usuarios, es conveniente usar más que uno, puesto que ningún buscador puede tener en su base de datos más que una parte limitada del total de los documentos disponibles en la WWW,

*estimándose que alrededor de la mitad de ellos no son encontrados por los motores de búsqueda durante sus visitas a la WWW.”*²⁰⁶

En otras palabras, los buscadores generales, son aquellos buscadores que ofrecen la posibilidad de localizar en su propia base de datos las direcciones que podrían responder aproximadamente a lo que queremos encontrar, siempre que se introduzcan los modificadores y datos adecuados.²⁰⁷

De otra parte, es de señalar que como función principal se le atribuye a los buscadores generales, la de alinear sobre la pantalla la mayor cantidad de direcciones Web teniendo presente la pregunta o palabra clave propuesta y la menor cantidad de tiempo en su procesamiento, de forma que esta se produzca casi al mismo tiempo en que se formula el tema del cual se desea obtener información²⁰⁸.

De acuerdo con lo anterior, es indiscutible que los buscadores generales ofrecen a los usuarios gran cantidad de información, sin embargo, como ya se mencionó, carecen de la selección rigurosa que debe presentarse en temas especializados, en este sentido, podemos afirmar que las listas que arrojan los resultados son en la totalidad de los casos inabarcables, si se tiene en cuenta las miles de referencias a las cuales se hace remisión para consulta, además de las páginas que carecen de precisión, no pertinencia, erróneas o que no cuentan con sustentos bibliográficos serios y que finalmente deben ser descartadas.

²⁰⁶ TEMPRANO SÁNCHEZ, A. (2008). *Diseño y desarrollo de un software para la creación de webquest*. Madrid: Bubok. p. 51.

²⁰⁷ MUÑOZ LÓPEZ, M. (2010). *Humanistas, textos e Internet: una difícil relación*. En el texto: *Humanismo y pervivencia del mundo clásico homenaje al profesor Antonio Prieto* (Vol. 4,). Alcañiz (Madrid) Editada por: J. M. Maestre Maestre, A. Prieto, J. P. Barea, & L. Charlo Brea, p. 1994.

²⁰⁸ BÉRA, M., & MECHOULAN, É. (1999). *La machine internet*. Paris: Ediciones Odile Jacob. p. 38.

En consecuencia, es importante aclarar que los buscadores generales no son la herramienta más apropiada para efectos de emprender búsquedas sobre temas jurídicos específicos, toda vez que con la gran cantidad de información contenida en la web, al arrojar los resultados de la búsqueda se puede obtener información inadecuada.

Ahora bien, dentro de las inconsistencias mencionadas con anterioridad, sobre los buscadores generales, encontramos la ardua tarea que debe realizar el usuario, al discriminar dentro de los listados de resultados la información que realmente considera pertinente para su objeto. Cabe anotar que, en el caso de temas especializados, como el derecho, los abogados podrían caer en imprecisiones o cometer errores a causa del uso de los buscadores genéricos, toda vez que dentro de los resultados habrá que emprender una lucha contra el ruido y el silencio informático.

En este sentido, el autor Manuel López Muñoz de la Universidad Almería, en su artículo *“Humanistas, textos e Internet: una difícil relación”*, menciona lo siguiente:

“el recurso a los buscadores tradicionales no es, precisamente, la mejor de las ideas cuando se necesita un material determinado en el menor tiempo posible. El problema es que las páginas especializadas tampoco son la mejor de las opciones, aun pareciéndolo prima facie. Hay toda una serie de problemas que aquejan a este tipo de recopilaciones, entre los que deben destacarse, de entrada, los derivados de los criterios de selección y mantenimiento de los enlaces. La selección es, por sí misma, el principal de

los escollos en cualquier trabajo, y no iba a ser menos; puede verse condicionada por la capacidad de búsqueda del editor (señalar lo que va encontrando, o lo que le gusta, o lo que encuentra después de escudriñar a conciencia), pero también por la propia naturaleza de Internet, en la que las direcciones electrónicas pueden aparecer y desaparecer con un ciclo lo suficientemente corto como para que no sea posible asegurar que nuestro repertorio de enlaces funciona. Y eso es, simple y llanamente, la desgracia suma de una página especializada, que se debate entre la necesidad de completar la información y la de verificar que esa información sigue existiendo. Se llega, así, a un proceso recursivo y retroalimentado que hace que, cuanta más información introducimos, más tiempo necesitamos para verificar su existencia, con lo que menos tiempo tenemos para buscar nuevos recursos, y así in acternum. Finalmente, el responsable de una página de este tipo debe optar entre buscar o mantener, a sabiendas de que cualquiera de las dos decisiones aboca al fracaso por falta de la otra. Son útiles, pues, las páginas especializadas, pero solo hasta cierto punto. Son la puerta de entrada pero no la solución definitiva.”²⁰⁹

En conclusión, en atención al concepto de buscador general, podemos afirmar que el uso de los mismos para usuarios de ciencias especializadas como el derecho, no resulta recomendable, si lo que se pretende es elaborar investigaciones serias sobre un tema puntual del ordenamiento jurídico, toda vez que al presentarse problemas como la inseguridad jurídica, bien se podría encontrar con información que ha

²⁰⁹ MUÑOZ LÓPEZ, M. (2010). op. cit. p. 1994.

perdido vigencia y utilizarla sin conocer esta situación. Sin embargo este es un tema que será estudiado dentro de los problemas que suscitan los buscadores.

2.8. Funcionamiento de los buscadores

En cuanto a lo relacionado con el funcionamiento de los *search engines* o buscadores, podemos mencionar que estos requieren de tres factores a saber: “1. La base de datos, 2. La interfaz con el usuario y, 3. La forma en que presentan los resultados de una búsqueda”²¹⁰ o el denominado motor de búsqueda que facilita la consulta a la base de datos.

En primer lugar, para la existencia de un buscador o *search engine*, se hace necesario crear una **base de datos** sólida integrada por los contenidos de la Web que se construyen gracias a los Robots o Web Crawler. En un principio esta se conformaba a partir de “la información proveniente de los formularios de alta que los autores o administradores de sitios les envían voluntariamente”²¹¹, no obstante, en la actualidad, el suministro de los documentos Web no se encuentra conformado únicamente con la información aportada de forma voluntaria, estos, a diferencia de los buscadores prístinos, cuentan con información extraída de diferentes fuentes, que incluso en algunas ocasiones es adquirida a partir de compras realizadas por el *search engine* o en el peor de los casos, a causa de la violación a la propiedad intelectual y en concreto a los derechos de autor.

²¹⁰ SIRI, L. (2000). op.cit. p. 59.

²¹¹ Ibídem. p. 60.

Sobre lo anteriormente mencionado, son múltiples las controversias generadas por el *motor de búsqueda*,²¹² GOOGLE a causa de la violación de la propiedad intelectual. Recientemente, el diario el Tiempo informó sobre la condena de 598.488 dólares, que recibió por haber colgado vídeos de autores de nacionalidad francesa, en su Web sin la previa autorización o aquiescencia de los mismos²¹³. Sin embargo, este caso no se encuentra aún concluido, toda vez que los abogados de GOOGLE recurrieron la decisión de primera instancia, alegando que en virtud de la aplicación de la ley en el espacio, se deben regir por el derecho francés y absolver teniendo en cuenta que este dispone, que los servidores no son responsables de los contenidos que cuelgan sus clientes.

Es importante mencionar, que en Francia se estableció al respecto, en jurisprudencia anterior, que los servidores de Internet únicamente se encuentran obligados a desaparecer los contenidos cuando sus autores así lo indiquen, de lo contrario estos no serían responsables. No obstante, cabe cuestionar si la aplicación de esta disposición se emplea igualmente para los motores de búsqueda.

²¹² Motor de búsqueda: es el sinónimo frecuentemente usado para hacer referencia a un buscador o browser, que es aquel sistema informático creado con el fin de encontrar archivos almacenados en la Web, los cuales se obtienen a partir de la introducción de *key words* o palabras claves.

²¹³“El Tribunal de Apelación de París condenó a Google por haber puesto en acceso libre a disposición del público fotografías de un artista francés, extractos de la película “Mondovino”, de un documental sobre el genocidio armenio y de otro sobre el caso de corrupción francés “Clearstream”. Según el gigante de internet estas denuncias son “muy antiguas”, anteriores a 2007, año en el que **Google** puso en marcha una serie de herramientas que permiten a los derechohabientes denunciar el acceso gratuito a sus contenidos a través de su plataforma y, posteriormente, eliminarlos. El Supremo francés estableció en una sentencia de febrero pasado que los servidores de internet sólo están obligados a quitar los contenidos cuando sus autores se lo indiquen, pero que no son responsables de los mismos.

Los denunciantes consideran que sus contenidos no desaparecieron de la página de Google después de que denunciaran su presencia sin derechos, por lo que consideran que el Tribunal de Casación (órgano supremo del sistema judicial francés) les dará la razón”. EL ESPECTADOR. *Google, condenado por colgar vídeos en su web sin permiso de los autores*. (2011, Marzo 9). Consultada en marzo 5, 2014. <http://www.elespectador.com/tecnologia/google-condenado-colgar-videos-su-web-sin-permiso-de-lo-articulo-255768>

Además de lo anterior, los buscadores o *search engines* cuentan con otra forma de recuperar información contenida en Internet, por medio de unos programas denominados, *spiders o robots*, que son mecanismos para recuperar información contenida en las páginas Web que se encuentran.

Los *spiders o robots*, son indexadores automáticos que buscan formar bases de datos íntegras, por medio de la búsqueda en la World Wide Web de una forma metódica. Otros términos que recibe son: *Ants*, *automatic indexers*, *bots*, *web spiders*, *web robots* o especialmente en la comunidad FOAF «friends of a friends»²¹⁴ se le denomina *soutters*.

Muchos sitios, en particular, *search engines*, usan la Web Crawler con el fin de actualizar datos, y crear una copia de todas las páginas Web, visitadas para procesar

²¹⁴ FOAF (an acronym of Friend of a friend) is a machine-readable ontology describing persons, their activities and their relations to other people and objects. Anyone can use FOAF to describe him or herself. FOAF allows groups of people to describe social networks without the need for a centralised database.

FOAF is a descriptive vocabulary expressed using the Resource Description Framework (RDF) and the Web Ontology Language (OWL). Computers may use these FOAF profiles to find, for example, all people living in Europe, or to list all people both you and a friend of yours know.[1][2] This is accomplished by defining relationships between people. Each profile has a unique identifier (such as the person's e-mail addresses, a Jabber ID, or a URI of the homepage or weblog of the person), which is used when defining these relationships.

The FOAF project, which defines and extends the vocabulary of a FOAF profile, was started in 2000 by Libby Miller and Dan Brickley. It can be considered the first Social Semantic Web application, in that it combines RDF technology with 'Social Web' concerns.

Tim Berners-Lee, in a 2007 essay,[3] redefined the Semantic web concept into something he calls the Giant Global Graph, where relationships transcend networks/documents. He considers the GGG to be on equal grounds with Internet and World Wide Web, stating that "I express my network in a FOAF file, and that is a start of the revolution."

Traducción propia: FOAF, (literalmente significa amigo de un amigo) Es un proyecto dentro de la Web semántica para describir relaciones mediante RDF que puedan ser procesadas. Esta es una aplicación de la Web semántica que es considerada como muy prometedora.

FOAF describe personas, sus actividades y sus relaciones con las otras personas y objetos. FOAF permite grupos de personas para describir las redes sociales sin la necesidad de centralizar la base de datos.

FOAF es un vocabulario descriptivo expresado usando RDF (Resource Description Framework) o Marco de Descripción de Recursos que es un Framework o Marco Estandarizado de de conceptos, prácticas y criterios, para enfocar un tipo de problemática particular.

Además FOAF es un vocabulario descriptivo que usa la (OWL o Web Ontology Language). Los computadores pueden usar FOAF por ejemplo para encontrar perfiles de toda la gente que vive en Europa o para encontrar amigos en común con un amigo que usted conoce. FOAF. (2015, Febrero 3). [http://en.wikipedia.org/w/index.php?title=FOAF_\(ontology\)&oldid=645417702](http://en.wikipedia.org/w/index.php?title=FOAF_(ontology)&oldid=645417702)

de forma posterior la información, por medio de un motor de búsqueda que indexa las páginas y proporciona un sistema de búsqueda eficaz y más íntegra. La función principal de la Web Crawler o los también denominados robots, consisten en emprender recorridos en la web y escrutarla, con el fin de conocer sus variaciones, nuevas creaciones web, para poder ofrecer la información actualizada.

El programa de Web Crawler,

“funciona visitando una lista de URLs, identificando los hiperenlaces en dichas páginas y añadiéndolos a la lista de URLs a visitar de manera recurrente de acuerdo con un determinado conjunto de reglas. La operación normal consiste en que se le da al programa un conjunto de direcciones iniciales, la araña descarga estas direcciones, analiza las páginas y busca enlaces a páginas nuevas. Luego descarga estas páginas nuevas, analiza sus enlaces y así sucesivamente.

Entre las tareas más comunes de las arañas de la Web tenemos:

- *Crear el índice de una máquina de búsqueda.*
- *Analizar los enlaces de un sitio para buscar links rotos.*
- *Recolectar información de un cierto tipo, como precios de productos para recopilar un catálogo”²¹⁵.*

²¹⁵ ARAÑA WEB. (2014, Octubre 21). Consultada en febrero 5, 2014.
http://es.wikipedia.org/w/index.php?title=Ara%C3%B1a_web&oldid=77670867.

En conclusión, podemos afirmar que las bases de datos se encuentran conformadas las más de las veces por información suministrada de forma voluntaria por los autores o administradores de determinado sitio Web, por la utilización de programas informáticos, específicamente creados para encontrar información nueva en Internet o por la extracción sin la aquiescencia de su autor que bien se puede realizar con el uso de programas informáticos antes mencionados: *Web crawlers, spiders o robots*.

Por otra parte, es necesario mencionar que el éxito del posicionamiento de una página web y su ubicación dentro de tal o cual base de datos, las más de las veces depende de la continua actualización del sitio Web, so pena de desaparecer dentro de los resultados de las búsquedas.

El segundo factor, incidente en el modo de trabajo de un search engine o un buscador, es ***la interfaz con el usuario***, que consiste en la comunicación entre el buscador y el usuario, para tales efectos, se hace necesario establecer aquello que el navegante tiene permitido hacer al ingresar palabras clave, es decir, es el puente de comunicación entre el buscador y el usuario, motivo por el cual las palabras claves o keywords, juegan un papel fundamental a la hora de emprender una búsqueda óptima.

Finalmente, tenemos que el último factor para el funcionamiento de los *search engines* o buscadores, es ***la presentación de resultados***, que juega un papel fundamental en la selección de la información requerida, para lo cual, se ofrecen resultados organizados teniendo en cuenta su jerarquía, popularidad o número de

visitas, pertinencia, orden alfabético, mayor coincidencia de la búsqueda en atención a las *keywords* introducidas, etc.²¹⁶.

A esa presentación de resultados, también se le ha atribuido la denominación de ***motor de búsqueda***, que es la página que vemos cuando realizamos la búsqueda, una vez introducidas las palabras claves, este es el encargado de cotejar con su base de datos la información solicitada y arrojar la información de forma organizada en atención a la relevancia de cada una de las páginas.

Una vez examinado lo anterior, tenemos que pese al cumplimiento de los tres anteriores factores, los buscadores en ningún caso pueden abarcar el 100% de la Web, motivo por el cual hay que tener presente que en cualquier búsqueda será determinante el uso del *search engine* indicado, teniendo en cuenta la naturaleza del tema y la calidad de los resultados que arrojen a cerca del mismo. Igualmente, es menester mencionar que en algunos casos, dentro de la organización de los resultados de la búsqueda se podrá observar en las primeras filas, páginas Web, con un alto rango de visitas, que no necesariamente pueden contener la información más relevante sobre la *keyword* introducida. De forma que siempre que se emprenda una búsqueda, lo idóneo será examinar las páginas subsiguientes a la arrojada como principal, esto con el fin de examinar contenidos con mejores resultados.

2.9. Límites de los buscadores

²¹⁶ “Existen otros criterios de jerarquización que pueden arrojar resultados inesperados. Por ejemplo, el algoritmo de AltaVista le da un puntaje más alto a las palabras únicas e inusuales”. SIRI, L. (2000). op.cit. p. 62.

En el presente acápite nos corresponde además estudiar, ¿Qué límite tienen los buscadores o *search engines* dentro de los resultados que pueden arrojar a partir de la introducción de *keywords*? Para estos efectos, podremos observar que los buscadores cuentan con ciertos criterios de organización y jerarquización de los resultados, de ahí que encontremos verbigracia: el modelo de *Page Rank* de GOOGLE, que no es otra cosa que la asignación que surge por conducto de una fórmula matemática, antes ya mencionada, que genera un guarismo determinando según su número de resultado, la posición dentro de la cual se deberá encuadrar en determinados resultados.

Pese a lo anterior, las fórmulas de los buscadores en general han sido drásticamente cuestionadas, en especial a causa de los resultados parcializados que estos manejan, toda vez que al realizar la labor de forma automatizada, se pretermiten criterios esenciales de conocimientos técnicos que se debe tener en cuenta al realizar dicha discriminación, por esta razón es que nosotros consideramos que las más de las veces, los buscadores no cumplen con los requisitos necesarios para suministrar resultados dentro de bases de datos especializadas, especialmente porque pueden llegar a incluir información no pertinente, ruidos informáticos, información errónea contenida en un sitio Web que hable sobre el tema objeto de búsqueda, entre otros problemas. Por lo anterior, bien se hará en usar esta herramienta tratándose de información de contenidos genéricos, que no especializados como los bancos de datos jurídicos.

Un límite con el cual cuenta la gran mayoría de los buscadores, es justamente que los *search engines* en su afán de obtener la más amplia lista de páginas web, arroja

resultados tan extensos que obligan al usuario a realizar un segundo trabajo de discriminación de información, para obtener los resultados verdaderamente queridos, por tal motivo observamos, que desde el punto de vista de la idoneidad, podríamos llegar a afirmar que los buscadores pretenden principalmente emitir la mayor cantidad de resultados teniendo en cuenta la coincidencia con las *Keywords* introducidas, suministrando contenidos que en gran variedad de casos, no guardan relación con la finalidad de la búsqueda.

Otro de los límites que presentan los buscadores, consiste en las dificultades de identificación del género, la temática, el autor, la fecha de creación, el país de origen, la institución, empresa u organización que lo emite y el lugar específico de su creación. Igualmente, la dificultad que presenta la realización de indexaciones, encontrar la longitud del texto y el tema.

Es notable observar, algunos defectos de las búsquedas realizadas por una máquina en contraposición a la selección y discriminación realizada por humanos, es así, como algunos sitios Web, pueden sesgar los procesos de selección con el fin de lograr su posicionamiento, además de pretermir en las búsquedas automáticas los análisis de objetivos, historia y políticas propias de un determinado sitio Web²¹⁷, hecho que genera inseguridad en el contenido de las búsquedas obtenidas.

Otro fundamento, que respalda la existencia de límites en las búsquedas automatizadas, consiste en la restricción que existe cuando se incluyen *Keywords* que no forman parte del contenido de una página Web que si contiene la información

²¹⁷ Ibidem. p. 68.

objeto de la búsqueda, pero que hace referencia a ella con la inclusión de otro vocabulario. Este problema persiste en la actualidad en los buscadores que existen y se presenta como una limitante de la búsqueda, de manera que hoy por hoy la sinonimia y la polisemia son aspectos que continúan causando controversia.

Como menciona The Clever Project, estos problemas podrían corregirse mediante la creación de ‘redes semánticas’ que almacenen información sobre las relaciones entre los diversos significados de las palabras. Pero este tipo de compilaciones corren el riesgo de solucionar el problema de la sinonimia agravando el de la polisemia²¹⁸.

Por otra parte, dentro de los resultados que arrojan los denominados *search engines*, encontramos diferentes formatos de archivo, es decir que además de textos, los buscadores arrojan videos, imágenes, audio, diferentes programas, formatos, motivo por el cual GOOGLE, ha adquirido a YOUTUBE²¹⁹, con el fin de poder arrojar búsquedas completas sobre la totalidad del contenido de la Web. No obstante, este es un tema que aún se encuentra en desarrollo y que busca integrar dentro de las búsquedas el máximo contenido de los diferentes formatos de la Web.

2.10. Search Engines mejor posicionados en la actualidad. Modelo De Google.

²¹⁸ Ibídem. p. 69. “The Clever Project desarrolló un sistema propio de búsquedas, el cual analiza las primeras 200 respuestas a una consulta efectuada en un buscador. Luego agrega a la lista todos los sitios que son enlazados desde esas páginas. Analiza estos últimos y selecciona sólo aquellos que tienen enlaces hacia el grupo original. Esto hace crecer el listado y a partir de allí, establece la relevancia según, entre otras cosas, la proximidad de las palabras clave utilizadas en la consulta realizada por el navegante con las usadas en los sitios que enlazan hacia los 200 originales. El objetivo es identificar aquellos que actúen como “catalogadores” de páginas web relacionadas con un tema determinado. Si hay varios “catalogadores” que apuntan hacia un mismo sitio, se supone que éste constituye una “autoridad” en la materia consultada, por lo que aquel que reciba más enlaces de los catalogadores ocupará la primera posición en dicha lista”.

²¹⁹ *Google compra YouTube* «1.650 millones de dólares». (2006, Octubre 9). Consultada en febrero 5, 2014. <http://google.dirson.com/post/2885-google-compra-youtube/>

En los primeros momentos del surgimiento de Internet, existían innumerables buscadores, sin embargo todos contenían falencias que afectaban los resultados, motivo por el cual no era recomendado realizar búsquedas especializadas en razón de su diseño, tal es el caso de WebCrawler, este buscador consideraba como unidades separadas el nombre de pila y el apellido, y no como una frase indivisible. Sin embargo, si uno buscaba un término excepcional, este buscador si era efectivo, dado que indexaba el texto completo de las páginas²²⁰.

En 1998, un acontecimiento va a marcar el antes y después de la era de los buscadores en internet, con el surgimiento de Google, se marca un hito en la historia, debido al nuevo enfoque de ordenamiento que atendió a criterios de relevancia, toda vez que en el año 2000 y siguientes, la gran preocupación y la estrategia principal de marketing consistía en el tamaño de las bases de datos, de manera pues, que se presenta una mutación de las necesidades, principalmente porque Google al contar con una amplia base de datos que abarca gran cantidad de los contenidos Web, soluciona este problema y le da paso a otro consistente en mejorar los criterios de relevancia con los cuales debe contar un buscador actual.

Con la llegada de Google al mercado de los buscadores y su posterior posicionamiento, varios buscadores desaparecieron, no obstante, paralelo al modelo de GOOGLE, subsisten los siguientes buscadores: Yahoo, Ask Altavista, MNS, AOL, AlltheWeb, Go, Netscape Search, Live, Lycos, Snap, Webcrawler, A9, blinkx, boing, bookmach.com, BOXXET, ChaCha, A9, AURA, blinkx, bookmach.com, ChaCha, ClipBlast!, Clusty, collarity, CometQ, CONGOO, d e c i p h o, del.icio.us, digg, Dmoz, exelead, factbites, fazzle, FEEDS|2.0, FindSounds, GIGABLAST,

²²⁰ SIRI, L. (2000). op.cit. p. 63.

girafa, gnn o d, GoDefy, goshme, Goyams buscador que suministra los contenidos en conjunto de GOOGLE, Yahoo y msn.

Igualmente, se encuentran presentes en la Web: ICEROCKET, search engine que se ha catalogado como confidencial, ixquick, KartOO, last.fm buscador especializado en música, Lexxealpha, likeque es un buscador recientemente adquirido por GOOGLE; propietario de una amplia base de datos de eCommerce desde el año 2006, LiveDeal expertos en marketing, Local.com Buscador local de estados unidos que ofrece información sobre Los Angeles, New York, Atlanta, Miami, Chicago, lurpo que es un directorio de la web, MetaGlossary que define a cerca de 2,000,000 de términos, frases and acrónimos, mnemomap, Mojeek, Mooter, MrSAPO que ofrece un multibuscador y una colección de links en línea organizado por diferentes temas.

En igual sentido encontramos buscadores como: nayio que suministra al usuario la posibilidad de introducir las keywords y encontrar los resultados de su búsqueda en idioma coreano, Octora, OiHoi Search, Ozu, pluggd el cual ofrece a sus usuarios un buscador enfocado en la búsqueda de videos, pronto.com buscador especializado en la adquisición de bienes, QTsearch buscador mexicano de tecnología, turismo o vacaciones, salud y belleza, Quintura, Releton, retrevo, gamma, SearchTheWeb2, SEEQPOD, sidekiq, Slifter; search engine dedicado a las compras, soople, Sphider, SPURL.net, S R C H R, SurfWax, Swoogle buscador de web semántica que cuenta con más de 10.000 ontologías, TagJag!, Terra, thefind.com, Trexy conocido como el buscador que combina los resultados de GOOGLE, Bing y Yahoo que realiza

recomendaciones basadas en lo que la gente busca actualmente, tiene su origen en Reino Unido desde el año 2003.

Para finalizar, encontramos en la web como *search engines*: turboscout, url.com, VMGO.com, Web 2.0 el cual ofrece una búsqueda simple por *keyword* relacionada con lo que el usuario desea trabajar con el fin de ver las herramientas de Web 2.0 y sus links disponibles, un ejemplo de links de búsqueda es: mashups, blogs, clipart, diseño, etiquetado social, AJAX aplicaciones, las herramientas para webmasters, calendarios, etc. Igualmente se encuentra presente en la Web: WEBBRAIN, whonu?, WIKIO, WiseNut buscador en idioma coreano, yoono, yoogle, YuFind, ZABASEARCH buscador de personas de origen estadounidense, zapmeta, ZUULA, Scirus, buscador científico.

En el mismo sentido, subsisten buscadores que contienen unas características específicas diferenciadoras de los demás *search engines*: AltaVista, HotBot, Excite, Lycos, que facilitan las restricciones de búsqueda teniendo en cuenta el idioma, las búsquedas sucesivas sobre temas indeterminados, compras, empleos, juegos, además de filtrar según el dominio. Otros tales como Northern Light, son buscadores más especializados que proporcionan portales de investigación estratégica de investigación de mercado, inteligencia competitiva, análisis de negocios, desarrollo de productos e investigación tecnológica.

No obstante lo anterior, pese a la existencia de la gran cantidad de *search engines* que ofrecen sus servicios en la Web, el buscador con mejor posicionamiento en la actualidad es GOOGLE, por una diferencia de un 89,94 %, seguido de BING con el

4,37% y de YAHOO con 3,93%, según un estudio realizado por la firma Stat Counter²²¹.

Por su parte, en Latinoamérica, los estudios que se han realizado sobre los buscadores muestran igualmente resultados que se acercan a los indicadores mundiales, hasta febrero de 2011, arrojan que en Colombia GOOGLE abarca la mayoría del mercado con un 95,21%, en segundo lugar BING con el 4,11% y en tercer lugar YAHOO con 0,36%. Lo propio sucede con países como Brasil y Chile quienes cuentan con mayores porcentajes en comparación con los colombianos.

Otro aspecto que es relevante mencionar a propósito de lo anterior, consiste en el crecimiento que ha tenido esta grande empresa desde sus inicios, si se tiene en cuenta el aumento considerable de empleados que forman parte de Google, toda vez que en el año 2003, contaba con 1.628 empleados, cantidad que pasó a ser 10.674 al final del año 2006, de forma que se incrementó cerca de 9000 empleados en menos de tres años. En junio de 2008, Google tuvo 19.604 empleados de tiempo completo, este hecho corresponde a causas financieras y a las excelentes condiciones laborales de sus empleados, quienes cuentan con el 20% de su tiempo laboral para trabajar en sus propios proyectos e investigaciones con plena libertad de elección. Motivo por el cual Google ha sido catalogado como una de las mejores empresas del mundo para trabajar²²².

²²¹ *Bing desplazó a Yahoo! y es el segundo buscador más usado en el mundo.* (2011, marzo 3). ENTER.CO: http://www.enter.co/cultura-digital/entretenimiento/bing-desplazo-a-yahoo-y-es-el-segundo-buscador-mas-usado-en-el-mundo/?s_cid=1

²²² LEVERING, R., & MOSKOWITZ, M. (n.d.). *100 best companies to work for*. Consultada en enero 12, 2012, Great Place to Work: <http://www.greatplacetowork.com/best-companies/100-best-companies-to-work-for>

De otra parte, en atención a la globalización Google ha logrado estar disponible en 130 idiomas, gracias a la contribución de personas dedicadas al tema, quienes han estado unidas intercambiando tips y trucos a cerca del mejor camino para traducir documentación. Aunque resulte increíble, muchas personas trabajan como voluntarias en una compañía de estados financieros absolutamente rentables, para ellos, la motivación va más allá de un salario, algunos se unen al proyecto de Google con fin de competir con otros programadores, otros por razones filantrópicas interesados por ser activistas, verbigracia: Existen traductores voluntarios. Estas comunidades de voluntarios juegan un papel fundamental en Google, al punto que sin ellos probablemente Google no tendría su marca tan posicionada como la tiene hoy en día.

Es importante considerar que, desde el año 2001, Google es conocida además de todo lo anterior, por haber comprado gran cantidad de compañías que contaban con productos interesantes, en promedio el gigante de la búsqueda ha adquirido cerca de 50 compañías. Algunas de ellas son: Blogging, trafic and map analysis, HTML, YouTube, Fledgling Company, entre otras. Conforme a esto, Google amplía sus servicios no solo de forma interna, gracias a la adquisición de reconocidas empresas. Para presentar una definición, bien podemos mencionar que

“Google es un motor de búsqueda que permite a los utilizadores encontrar sitios en la web, por medio de palabras claves. Es suficiente ingresar una o varias palabras en el campo de búsqueda que se estime están contenidas en el sitio que se busca y Google atraviesa la web, mediante el uso de robots informáticos, para la búsqueda de un contenido correspondiente a la

pregunta. Aparece entonces, sobre la pantalla un cierto número de sitios identificados por un título, algunas palabras y una dirección URL (por ejemplo www.xyz.be) sobre la cual es suficiente hacer un clic para dirigirse automáticamente al sitio concerniente."²²³

Una vez examinadas las cifras anteriores y su definición, cabe cuestionarnos sobre ¿Cuál ha sido el éxito que logró posicionar a GOOGLE como el buscador más usado en el mundo? uno de los motivos por el cual Google, es actualmente el buscador más usado y con mayor prestigio consiste en la fórmula registrada y patentada del método PageRank²²⁴, creado por sus cofundadores Larry Page y Sergey Brin.

²²³ R.G. 2007/AR/1730. R n° 2011/2999. N°: 817. 5 – 05 – 2011. LA COUR D'APPEL DE BRUSELAS, 9 CHAMBRE. En causa de: Google INC, sociedad de derecho americano ubicada en Mountain View, 1600 amphitheatre parkway, 94043 California (USA). Contra: Copiepresse; sociedad civil constituida en forma de una sociedad cooperativa de responsabilidad limitada donde su sede social es establecida en 1070 bruseles, boulevard paepsem, Societe de Droit d'Auteur des Journalistes (abreviación SAJ). Sociedad de Derechos de Autor y de los Periodistas y Assucopie.

²²⁴ "PageRank es una marca registrada y patentada por Google el 9 de enero de 1999 que ampara una familia de algoritmos utilizados para asignar de forma numérica la relevancia de los documentos (o páginas web) indexados por un motor de búsqueda. Sus propiedades son muy discutidas por los expertos en optimización de motores de búsqueda. El sistema PageRank era utilizado por el popular motor de búsqueda Google para ayudarlo a determinar la importancia o relevancia de una página. Fue desarrollado por los fundadores de Google, Larry Page y Sergey Brin, en la Universidad de Stanford".

PageRank confía en la naturaleza democrática de la web utilizando su vasta estructura de enlaces como un indicador del valor de una página en concreto. Google interpreta un enlace de una página **A** a una página **B** como un voto, de la página **A**, para la página **B**. Pero Google mira más allá del volumen de votos, o enlaces que una página recibe; también analiza la página que emite el voto. Los votos emitidos por las páginas consideradas "importantes", es decir con un PageRank elevado, valen más, y ayudan a hacer a otras páginas "importantes". Por lo tanto, el PageRank de una página refleja la importancia de la misma en Internet".

El algoritmo inicial del PageRank lo podemos encontrar en el documento original donde sus creadores presentaron el prototipo de Google: "The Anatomy of a Large-Scale Hypertextual Web Search Engine".

$$PR(A) = (1 - d) + d * \sum_{i=1}^n \frac{PR(i)}{C(i)}$$

Donde:

- **PR(A)** es el PageRank de la página A.
- **d** es un factor de amortiguación que tiene un valor entre 0 y 1.
- **PR (i)** son los valores de PageRank que tienen cada una de las páginas **i** que enlazan a A.
- **C(i)** es el número total de enlaces salientes de la página **i** (sean o no hacia A).

Algunos expertos aseguran que el valor de la variable **d** suele ser 0,85. Representa la probabilidad de que un navegante continúe pulsando links al navegar por Internet en vez de escribir una url

En aras de proporcionar prioritariamente a los usuarios, los mejores resultados, de forma organizada, en 1996, se cuestionó la practicidad de realizar revisiones de listas prolongadas en los resultados arrojados en la búsqueda, sin discriminar de forma alguna la información, por tal motivo en el seno de la Universidad de Stanford, los estudiantes Larry Page y Sergey Brin, presentaron como tema de tesis la clasificación de los resultados de los buscadores o *search engine* en la Web. Estos dos estudiantes de posgrado de la Universidad de Stanford, desarrollaron un motor de búsqueda denominado "BackRub" que utilizaba enlaces para determinar la importancia de cada página web. En 1998, ya habían formalizado su trabajo y creado la compañía que actualmente se conoce como Google²²⁵.

directamente en la barra de direcciones o pulsar uno de sus marcadores y es un valor establecido por Google. Por lo tanto, la probabilidad de que el usuario deje de pulsar links y navegue directamente a otra web aleatoria es $1-d$. La introducción del factor de amortiguación en la fórmula resta algo de peso a todas las páginas de Internet y consigue que las páginas que no tienen enlaces a ninguna otra página no salgan especialmente beneficiadas. Si un usuario aterriza en una página sin enlaces, lo que hará será navegar a cualquier otra página aleatoriamente, lo que equivale a suponer que *una página sin enlaces salientes tiene enlaces a todas las páginas de Internet*.

El peso o importancia de una página es el resultado de una "votación" entre todas las demás páginas de la World Wide Web acerca del nivel de importancia que tiene esa página. Un hiperenlace a una página cuenta como un voto de apoyo. El PageRank de una página se define **recursivamente** y depende del número y PageRank de todas las páginas que la enlazan. Una página que está enlazada por muchas páginas con un PageRank alto consigue también un PageRank alto. Si no hay enlaces a una página web, no hay apoyo a esa página específica. El PageRank de la barra de Google va de 0 a 10. Diez es el máximo PageRank posible y son muy pocos los sitios que gozan de esta calificación, 1 es la calificación mínima que recibe un sitio normal, y cero significa que el sitio ha sido penalizado o aún no ha recibido una calificación de PageRank. Parece ser una escala logarítmica. Los detalles exactos de esta escala son desconocidos.

PageRank ha tomado su modelo del *Science Citation Index* (SCI) elaborado por Eugene Garfield para el Instituto para la Información Científica (ISI) en los Estados Unidos durante la década del 50. El SCI pretende resolver la asignación objetiva de méritos científicos suponiendo que los investigadores cuyo factor de impacto (número de publicaciones y/o referencias bibliográficas en otros trabajos científicos) es más alto, colaboran en mayor medida con el desarrollo de su área de investigación. El índice de citación es un elemento determinante para seleccionar qué investigadores reciben becas y recursos de investigación. PAGE RANK. (2015, enero 27). Consultada febrero 15, 2015. <http://es.wikipedia.org/w/index.php?title=PageRank&oldid=79638444>.

²²⁵ GOOGLE EMPRESA. (n.d.). *Nuestra historia en Profundidad*. Consultada en mayo 9, 2015. <https://www.google.com/about/company/history/?hl=es>

Esta fórmula, es un valor numérico que representa la acogida o importancia con la cual cuenta una página Web en internet, con el fin de organizar la información en términos democráticos. Esta original fórmula matemática, examina cual debe ser el posicionamiento de la página, teniendo en cuenta un margen de error que va entre 0 y 1, al cual se le suman los valores de *PageRank* que tienen como resultado todas y cada una de las páginas que se enlazan a la página objeto de análisis, dividido sobre el número de enlaces salientes de cada página que enlazó la analizada, esta vez sin tener en cuenta si este es o no hacia la página objeto de estudio.

La fórmula del *PageRank*, propuesta por Larry Page, surgió gracias a la analogía que realizó de las páginas Web en semejanza con los autores de artículos científicos, quienes miden su acogida dentro del mundo académico a partir de la cantidad de veces que ha sido citado por otros autores, teniendo en cuenta a su vez la trayectoria y el reconocimiento de estos últimos dentro del ámbito académico, de esta manera se alcanza el reconocimiento y la aceptación de sus teorías o propuestas. En ese orden de ideas, según la fórmula de *PageRank*, “la importancia de la información de una página web estaba íntimamente relacionada con la cantidad –y calidad– de páginas que la enlazaban”²²⁶ por lo cual estos enlaces son equivalentes a una citación dentro del mundo académico.

Este planteamiento robusto de una lógica pragmática, otorgaron la posibilidad a su socio Sergey Brin,²²⁷ de materializar y poder determinar por intermedio de una

²²⁶ *Page Rank, la idea de Larry que cambió el mundo*. (2011, abril 2). Consultada en febrero 5, 2014. ENTER.CO: <http://www.enter.co/cultura-digital/redes-sociales/pagerank-la-idea-de-larry-que-cambio-al-mundo/>

²²⁷ Es menester hacer claridad que los creadores de GOOGLE, Larry Page y Sergey Brin, contaron con la colaboración de la Universidad de Stanford en 1995, que posee un ambiente favorable en términos de desarrollo de tecnología web, en comparación con los más altos estándares de

fórmula matemática compleja la popularidad y reconocimiento de las páginas Web. Sin embargo, es importante mencionar que pese a que la fórmula del *PageRank*, se encuentra citada en líneas anteriores, esta conserva algunos enigmas, que muy seguramente buscan proteger el *know how* del buscador que cuenta con mayor popularidad en el mundo: GOOGLE.

Una vez puesto en marcha, la idea de Larry, se inicia un proyecto denominado BackRub, encargado de analizar cada uno de los 10 millones de documentos que reposaban en la Web con el fin de establecer los enlaces con que contaba cada página. Con los resultados de BackRub, se pudo llegar a la conclusión que la sumatoria de los enlaces Web que conducían a una página en particular permitían otorgarle una calificación o ranking a la página. “No obstante, para que la organización fuera precisa, era necesario que los enlaces estuvieran valorados de acuerdo con el ranking que tenía la página donde fue publicado cada uno (en otras palabras, los enlaces o recomendaciones tenían mayor valor si provenían de sitios muy enlazados o recomendados)”²²⁸.

De otra parte, Google en lugar de realizar el posicionamiento de una página web, en atención a la cantidad de visitas y tiempo transcurrido en la consulta de la página, examina la naturaleza de la producción científica, la calidad de las citas, su importancia dentro del tema que se analiza y las veces que ha sido citado o se ha hecho referencia a propósito de este en otros artículos.

universidades como: University of California o Berkeley. Cuando Page y Brin explicaban la misión de Google, mencionaban que era principalmente organizar el mundo de la información y hacerla universalmente accesible y útil.

²²⁸ *Page Rank, la idea de Larry que cambió el mundo*. (2011, abril 2).op. cit.

Como hicimos referencia, con anterioridad, en Google cada página web es asimilada a los artículos científicos que gozan de reconocimiento gracias a la cantidad de citas que se referencian en otras páginas web. Si se incluye un enlace dentro de una página web, lo que con ello se pretende es recomendar la existencia de otra página a propósito del mismo asunto. De manera que al contar la cantidad de citaciones, de forma automática se sabrá la ubicación dentro de los resultados de la búsqueda. Es importante mencionar que esta es la base del algoritmo de Google.

Pese a la explicación anterior, el algoritmo de Google no solo cuenta la cantidad de citaciones que se realizan sobre un artículo en la web, este buscador además discrimina cada uno de los artículos en atención a su relevancia, verbigracia: es diferente encontrar una cita de un artículo proveniente de una revista indexada de una prestigiosa universidad, que un artículo escrito en un blog informal de la web. Así cómo funcionan los artículos dentro del ámbito académico, funcionan de igual forma las páginas que tienen mayor cantidad de citaciones y las páginas que cuentan con pocos enlaces en otras páginas web.

De otra parte, además de la mencionada fórmula del page Rank, Google cuenta con otros mecanismos que restringen y proporcionan mejores resultados en la búsqueda, entre ellos cabe mencionar las sencillas fórmulas que aparentemente podrían parecer insignificantes, pero que a la hora de emprender una búsqueda exitosa pueden marcar una amplia diferencia, la primera de ellas, es aquella que mide y le otorga más importancia a la mayor proximidad entre las palabras clave y el sistema que otorga más valor a los enlaces de sitios con muchos enlaces entrantes, pero pocos enlaces salientes.

Es menester mencionar, que los anteriores mecanismos que restringen la búsqueda, no se pueden utilizar en atención a las preferencias del usuario, sino que Google lo hace, de forma automática, verbigracia: En lo relacionado con el mecanismo que le otorga mayor proximidad a las palabras clave, el usuario no puede elegir la proximidad de las palabras a diferencia de otros buscadores que si lo permitían.

En igual sentido, otro de los mecanismos que utiliza Google para la recuperación de la información consiste en el uso de la memoria “caché”, mención que aparece sobre los resultados de la búsqueda, de forma que cuando Google explora la web por conducto de un webcrawler o spider, “realiza una copia de cada página examinada y la almacena en una memoria caché, la cual permite consultar esta copia en todo momento, y en particular en el caso donde la página original (o Internet) fuera inaccesible.”²²⁹

Cuando un usuario hace clic sobre el link <copia caché> de una página web, “Google la muestra en el estado en el cual ella se encontraba cuando se hizo la indexación más reciente. Por otra parte, el contenido caché es aquel sobre el cual se basa Google para determinar si una página es pertinente para sus consultas. Cuando una página caché es publicada, ella está precedida de un encabezamiento encuadrado que recuerda que allí se encuentra la copia caché de la página y no la página original, y cita los términos de la consulta que hayan encuadrado su inclusión en los

²²⁹ R.G. 2007/AR/1730. R n° 2011/2999. N°: 817. 5 – 05 – 2011. LA COUR D’APPEL DE BRUSELAS, 9 CHAMBRE. En causa de: Google INC, sociedad de derecho americano ubicada en Mountain View, 1600 amphitheatre parkway, 94043 California (USA). Contra: Copiepresse; sociedad civil constituida en forma de una sociedad cooperativa de responsabilidad limitada donde su sede social es establecida en 1070 bruselas, boulevard paepsem, Societe de Droit d’Auteur des Journalistes (abreviación SAJ). Sociedad de Derechos de Autor y de los Periodistas y Assucopie.

resultados de la búsqueda. Para facilitar la explotación de esta página, los diferentes términos de búsqueda son igualmente subrayados en colores diferentes.”²³⁰

“El link <<en caché>> dirige al internauta a la copia archivada de la página registrada. La advertencia contenida en la banda superior es redactada como a continuación:

<<Aquí la versión de Google de la página introducida en caché de http:// [...] extraída de la página de [...] GMT.

La versión <<en caché>> propuesta por Google corresponde a la página tal cual se presenta cuando se efectuó la última consulta por Google.

Se puede que la página haya estado modificada después de esa fecha. Haga clic aquí para obtener únicamente el texto incluido en caché.

Para crear un link con esta página o incluir en sus favoritos utilice la dirección siguiente [...]

Los siguientes términos de la búsqueda se han destacado [...].”²³¹

“Cuando la página <<en caché>> aparece sobre la pantalla del internauta, los elementos gráficos que no conciernen a lo propiamente hablado en el texto de búsqueda original (como por ejemplo los renvíos o hipervínculos a otros artículos, el

²³⁰ Ibídem.

²³¹ Ibídem.

estado del clima, las bolsas, los banners publicitarios eventuales) no son los que estaban fijados cuando se hizo el primer referenciamiento de esta página, pero estos son insertados, en tiempo real, por el servidor del sitio referenciado. Google no conserva en su memoria <<caché>> sino los textos extraídos de la página, convertidos en lenguaje HTML (Hiptertext Markup Language que es un formato de datos diseñado para representar las páginas web).

Cuando el internauta hace clic sobre el link <<en caché>>, él es dirigido a una copia archivada en la página web, registrada en Google en vez del lugar del sitio web de origen de la página. Pero como precisamos más arriba, otros elementos gráficos diferentes al texto son enviados por el servidor del editor. Por esto no es extraño, que el texto buscado sea fechado de muchos días anteriores que preceden a las otras informaciones que aparecen en la misma página porque el primero emana de la memoria <<caché>> de Google, mientras que los segundos provienen del servidor del editor.”²³²

La memoria caché, tal y como queda dilucidado en el párrafo anterior, no es entendida como parte integrante y esencial de un proceso técnico donde su única finalidad es permitir la transmisión en una red entre terceros por un intermediario; o una utilización lícita de una obra protegida, y que no tiene significación económica independiente, sino que esta, se presenta como un acto material de reproducción de páginas publicadas por un editor, que a todas luces viola la protección contenida en la Ley de Derecho de Autor LDA artículo 1, que establece que:

²³² *Ibíd.*

El autor es el único que tiene derecho de reproducir su obra o de autorizar la reproducción, de cualquier manera o sobre cualquier forma que sea, directa o indirectamente, provisional o permanente, en todo o en parte. Al igual, solo el autor tiene derecho de comunicarla al público por cualquier medio, este comprende la puesta en disposición del público de manera que cada uno pueda tener acceso en el lugar y momento que elija individualmente.²³³

Otro aspecto relevante a la hora de hablar del modelo de Google, consiste en la combinación que realizan de la teoría de la programación y la sociología de la red²³⁴, esto que expuesto con palabras sencillas, de fácil comprensión, contiene un conjunto de complejas fórmulas matemáticas enigmáticas que logran cumplir lo planteado en

²³³ Loi relative au droit d'auteur et aux droits voisins (LDA) 30 juin 1994. Publication: 27-07-1994
numéro: 1994009586. page: 19297

Dossier numéro: 1994-06-30/35. Entrée en vigueur: 01-08-1994 CHAPITRE I. - Du droit d'auteur.

Section 1. - Du droit d'auteur en général. Article 1. § 1. L'auteur d'une oeuvre littéraire ou artistique a seul le droit de la reproduire ou d'en autoriser la reproduction, de quelque manière et sous quelque forme que ce soit (, qu'elle soit directe ou indirecte, provisoire ou permanente, en tout ou en partie). <L 2005-05-22/33, art. 1, 005; En vigueur: 27-05-2005>

Ce droit comporte notamment le droit exclusif d'en autoriser l'adaptation ou la traduction.

Ce droit comprend également le droit exclusif d'en autoriser la location ou le prêt.

L'auteur d'une oeuvre littéraire ou artistique a seul le droit de la communiquer au public par un procédé quelconque (, y compris par la mise à disposition du public de manière que chacun puisse y avoir accès de l'endroit et au moment qu'il choisit individuellement). <L 2005-05-22/33, art. 1, 005; En vigueur: 27-05-2005> (L'auteur d'une oeuvre littéraire ou artistique a seul le droit d'autoriser la distribution au public, par la vente ou autrement, de l'original de son oeuvre ou de copies de celle-ci. La première vente ou premier autre transfert de propriété de l'original ou d'une copie d'une oeuvre littéraire ou artistique dans la Communauté européenne par l'auteur ou avec son consentement, épuise le droit de distribution de cet original ou cette copie dans la Communauté européenne.) <L 2005-05-22/33, art. 1, 005; En vigueur: 27-05-2005> § 2. L'auteur d'une oeuvre littéraire ou artistique jouit sur celle-ci d'un droit moral inaliénable. La renonciation globale à l'exercice futur de ce droit est nulle. Celui-ci comporte le droit de divulguer l'oeuvre. Les oeuvres non divulguées sont insaisissables. L'auteur a le droit de revendiquer ou de refuser la paternité de l'oeuvre. Il dispose du droit au respect de son oeuvre lui permettant de s'opposer à toute modification de celle-ci. Nonobstant toute renonciation, il conserve le droit de s'opposer à toute déformation, mutilation ou autre modification de cette oeuvre ou à toute autre atteinte à la même oeuvre, préjudiciables à son honneur ou à sa réputation.

²³⁴ GIRARD, B. op. cit. p. 16.

la teoría, es por este motivo que se guarda absoluta reserva sobre las mismas, en vista que este es el *know how* del motor de búsqueda más utilizado en el mundo.

Las fórmulas matemáticas están en todas partes de Google, los líderes de este buscador, ponen el poder de la computación al servicio del manejo matemático, tanto que una de las razones de su reconocido éxito, consiste en la creación de fórmulas como la del Page Rank, pocas compañías manejan tanta cantidad de matemáticas en sus relaciones y vocabulario.

Los ingenieros de Google usan fórmulas matemáticas para ver la velocidad y el contorno de un problema. Su cultura matemática conduce a buscar el principio general del comportamiento del usuario individual. Esta cultura matemática permite a los ingenieros no solo aplicar soluciones, también interesarse en la comunidad científica y en movilizar las búsquedas con el fin de encontrar soluciones a los diferentes problemas.

Visto lo anterior, cabe señalar que Google además de ofrecer búsquedas generales, ofrece varios servicios específicos, tal es el caso de Google News, que surgió a partir de la idea de Krishna Bharat, una ingeniera india, que utilizó el 20% del tiempo ofrecido por Google, para trabajar en un proyecto que nació gracias a la fascinación que tenía por ver a su abuelo en sintonía con la BBC todos los días con el fin de comparar los comentarios británicos con aquello que había leído en la prensa India²³⁵.

²³⁵ Ibídem. p. 79.

Por otra parte, Google, al propender por el suministro de información seria y fidedigna, se ha vinculado con la comunidad académica, en igual sentido, la estimulación a sus trabajadores hacia la investigación, genera que todas aquellas preguntas o temas que surgen como inquietudes del manejo, se puedan materializar en fórmulas matemáticas aplicadas a las diferentes herramientas de este buscador, con el fin de innovar y ofrecer la mayor cantidad de opciones de búsqueda a los usuarios.

Al igual que la idea anterior, Google's Desktop surgió a partir de la conversación entre amigos en la cual mencionaron a un ingeniero australiano que había creado un buscador para encontrar archivos locales en su computador Linux. Otro producto de Google, como Google Local, tiene origen en un concurso de programación convocado por Google, dentro del cual Daniel Egnor fue el ganador, con la creación de una herramienta de búsqueda que muestra únicamente resultados locales, este servicio compite con las páginas amarillas y con los servicios locales en Estados Unidos²³⁶.

Google también ofrece desde 2003, un servicio que consiste en los Blogs, desde que adquirió Blogger, incluso cientos de empleados han hecho uso de esta herramienta para crear su propio blog, de manera que, con la ayuda de la web social esta plataforma permite a los usuarios exponer sus ideas, planteamientos, innovaciones, a su vez analizar la acogida de la misma y estudiar los diferentes puntos de vista de sus participantes. Además de las anteriores ventajas, los beneficios de los blogs para negocios como Google son altamente efectivos, tal es el caso de los errores que se

²³⁶ Ibídem. p. 82.

pueden corregir gracias a la interacción con los usuarios de los servicios, toda vez que la información se encuentra concentrada.

En este mismo sentido, Google también incluye a los usuarios en el desarrollo de los procesos, gracias al uso de la Web 2.0 y más exactamente de los blogs, los usuarios son el soporte para mejoramiento e implementación de las nuevas herramientas. Esto habilita a la compañía para dividir los usuarios en dos segmentos: Aquellos que adoptaron los recursos de forma temprana y los usuarios comunes. Los primeros, que son usualmente mejor calificados y más interesados con la compañía, tienden a ser tolerantes con el producto porque entienden que al ser nuevo puede tener errores. Ellos tratan de evaluar y mejorar el proceso. Por otro lado los usuarios comunes tienden a ser más cautos y necesitan algún tiempo para familiarizarse con las actualizaciones.

Lo anterior, tiene por objetivo principalmente identificar los errores e implementar los arreglos o mejoras, por ejemplo: Google Book, en 2006 tenía un error en cuanto a su contenido, solo con los usuarios iniciales este se pudo detectar. Con el tiempo, más personas descubrieron el servicio y el error fue solucionado.

Ahora bien, una de las grandes ventajas de Google, en contraposición con otros buscadores, consiste en la gratuidad de sus servicios, lo cual creó un ambiente propicio para el posicionamiento de esta gran empresa en el mercado, toda vez que la sostenibilidad de los buscadores se basaba en el cobro de una suscripción que otorgaba el derecho a acceder a la información.

En este punto cabe cuestionarse entonces sobre ¿Cómo una empresa que ofrece sus servicios de forma gratuita, ha llegado a ser parte de las más rentables empresas del mundo financieramente hablando? Para resolver este interrogante, podemos sostener que el servicio principal de búsqueda que ofrece Google, constituye el medio de captación de los clientes que generan sus verdaderos ingresos. En términos de Marcel Mauss's, este modelo económico sería el denominado Potlatch²³⁷, que en palabras sencillas consiste en dar más para recibir más, de esta manera Google otorga la posibilidad a los usuarios de acceder a la información de forma gratuita, con el fin de hacer uso de la misma, copiar e incluso modificar sus contenidos.

En contraposición, los usuarios contribuyen al mejoramiento de los resultados de las búsquedas, verbigracia: Con la creación de nuevos contenidos web, tales como páginas en formato HTML, XML, Java, entre otros, que incrementan la información contenida en Internet, para que con posterioridad los Web Crawler identifiquen su presencia en la Web, sea incluido dentro de las bases de datos de Google y proporcionar así la integralidad de la información que se aporta al usuario.

Examinado lo anterior y con el fin de responder al planteamiento expuesto en líneas anteriores, la fuente de ingresos de Google, tiene explicación de igual forma, en la plataforma de publicidad que ofrece a sus usuarios, toda vez que al tener un acceso gratuito la cantidad de usuarios es ostensible, no solo por ofrecer un acceso libre, dado que existen compañías como Yahoo!, Excite o Bing que también ofrecen el mismo acceso, sino por el *Know How* con el cual cuenta este buscador; esto es: La fórmula de PageRank y las demás antes ya estudiadas, que crean una atracción

²³⁷ HERNÁNDEZ GÓMEZ, J. R. (2007). *Antropología jurídica Marcel Mauss el derecho del potlatch*. Bogotá. p. 47.

mucho mayor de usuarios de forma que en términos publicitarios es un buen medio para promocionar los diferentes productos.

Para finalizar, el análisis del financiamiento de Google, es necesario exponer una fuente de ingreso que contraria a las anteriores, afecta la confiabilidad de los resultados de la búsqueda, esta consiste en las ofertas que se presentan a los anunciantes para lograr un mejor posicionamiento de un sitio web y lograr estar en el top de los resultados²³⁸. Lo anterior constituye una de las críticas atribuidas a este gran buscador.

En este mismo sentido e infortunadamente, de forma paralela a la existencia del PageRank y de todas las fórmulas matemáticas que hacen de Google una herramienta que facilita la búsqueda general en la web, existen métodos para mutar los resultados, tales como la existencia de hipervínculos de otras páginas web, en lugares como foros, blogs, diccionarios, redes sociales, etc. Esto con el fin de incrementar las visitas producidas no solamente mediante el buscador, sino también mediante enlaces que conducen a la página que se pretende posicionar. A esto se le ha denominado spam²³⁹, no obstante google, en su afán por combatir las alteraciones de su fórmula de PageRank, cuando calcula el peso de una página, hace un esfuerzo por no tomar en cuenta los links que tengan el atributo de spam.

Otro de los problemas, que aqueja no solo al modelo de Google, sino a todos los buscadores que continúan en funcionamiento, consiste en que al pretender realizar una clasificación, cuando se arrojan los resultados de la búsqueda suministran la

²³⁸ GIRARD, B. op. cit. p. 30.

²³⁹ Se denomina *Spam*, a los mensajes no solicitados de tipo publicitarios, que se publican de forma masiva.

información de acuerdo con su nivel de relevancia, de forma tal que aparecen los resultados con mejor posicionamiento en los primeros lugares, esto trae como consecuencia la parcialización del resultado de la búsqueda, toda vez que no siempre las primeras páginas son aquellas que contienen la información requerida, en especial por los problemas de variación que se presenta con fórmulas como el *PageRank*.

Lo anterior se explica en mayor medida, dado que en atención a criterios como los del *PageRank*, dentro de los cuales para determinar la importancia de una página, se aplica la fórmula que determina la popularidad y cantidad de personas que ingresan, mas no la utilidad, pertinencia y relevancia de la información, motivo por el cual los buscadores de este tipo, en lo que tiene que ver con informática jurídica, no cumplen los criterios para proporcionar información idónea a los usuarios.

Otro de los problemas que se presenta con el *PageRank*, es justamente que su fórmula puede ser variada, por lo cual muchos buscan pagar para posicionar su página²⁴⁰, esto sin olvidar la existencia del *spam*, que modula y altera los resultados que surgen a partir de la aplicación del *PageRank*. En la actualidad, los posicionadores de buscadores han adquirido una notable importancia, dentro del ámbito jurídico “se ha desarrollado un cuerpo entero de jurisprudencia alrededor de cuestiones que afectan al posicionamiento en los buscadores, como el uso de marcas

²⁴⁰ “comScore estima que el 40-45 por ciento de las búsquedas incluyen los resultados patrocinados”. FALLOWS, D., LEE, R., & GRAHAM, M. op. cit.

registradas y metatags. La venta de posiciones en algunos buscadores ha creado también controversia entre bibliotecarios y defensores de los consumidores”²⁴¹.

Por otra parte, otra de las críticas que no solo se le puede atribuir a Google sino a los demás buscadores, consiste en la trascendencia de los límites que estos deben tener presentes, siempre que se trate de protección de datos personales y derecho de propiedad intelectual, esto en razón del papel preponderante que deben jugar estos derechos no solo en la industria de Internet, sino también en todas las relaciones sociales.

En torno a esto, cabe mencionar que varias han sido las controversias jurídicas presentes en Google, a causa de publicar contenidos protegidos por los derechos de autor. Es importante tener en cuenta que para el derecho europeo la protección de datos personales es un asunto de la dignidad humana, mientras que para Estados Unidos y países como Suecia, el punto de atención se centra en la protección del acceso público a la información y la transparencia. En mi sentir, la protección de datos personales, es especial si se trata de datos sensibles, es imprescindible, dado que de no existir, la intimidad de las personas y los demás derechos fundamentales quedarían sin tutela. No obstante, con la anterior afirmación no pretendemos en ningún momento ocultar la necesidad que tiene la sociedad de poder acceder a la información, que por su naturaleza deberá ser pública y de acceso libre.

²⁴¹ “En 2006, la posición en los buscadores es más importante que nunca, tanto que la industria ha desarrollado ("posicionadores en buscadores") para ayudar a los desarrolladores web a mejorar su posición en el buscador, y se ha desarrollado un cuerpo entero de jurisprudencia alrededor de cuestiones que afectan al posicionamiento en los buscadores, como el uso de marcas registradas en metatags. La venta de posiciones en buscadores por algunos buscadores ha creado también controversia entre bibliotecarios y defensores de los consumidores”. *Historia de Internet*. (2015, febrero 14). Consultada en febrero 5, 2013.

Wikipedia: http://es.wikipedia.org/w/index.php?title=Historia_de_Internet&oldid=80016630

Uno de los varios litigios que ha suscitado pronunciamientos jurídicos e interminables debates, es la reciente sentencia dentro de la cual un grupo que representa varios diarios belgas presenta una demanda en contra de Google por violación de derechos de autor²⁴². En este caso la Corte de Bélgica profirió un fallo a favor de Copiepresse y otras dos sociedades protectoras de los derechos de autor, donde reconoció explícitamente que Google había infringido las normas relativas a la protección de los autores de artículos de prensa, publicados en diarios belgas.

Uno de los más significativos problemas de Google, que se menciona dentro de la parte motiva de la citada sentencia, consiste en que la memoria caché de este motor de búsqueda²⁴³, contraviene las disposiciones relativas a la transitoriedad de la

²⁴² R.G. 2007/AR/1730. R n° 2011/2999. N°: 817. 5 – 05 – 2011. LA COUR D'APPEL DE BRUSELAS, 9 CHAMBRE. En causa de: Google INC, sociedad de derecho americano ubicada en Mountain View, 1600 amphitheatre parkway, 94043 California (USA). Contra: Copiepresse; sociedad civil constituida en forma de una sociedad cooperativa de responsabilidad limitada donde su sede social es establecida en 1070 bruselas, boulevard paepsem, Societe de Droit d'Auteur des Journalistes (abreviación SAJ). Sociedad de Derechos de Autor y de los Periodistas y Assucopie.

²⁴³ *Ibidem*. “Es común que Google registre en sus servidores una copia de las páginas que son regularmente visitadas por sus robots informáticos cuando realizan el referenciamiento de aquellas en el cuadro de servicio <<Google Web>>. Está igualmente establecido que cuando el internauta hace clic sobre el link <<en caché>>, Google le transmite esta copia.

Para decidir si ha habido reproducción y comunicación al público por parte de Google, importa poco que los elementos gráficos anexos (publicidad, meteorología, bolsa y reenvíos a otros artículos) sean transmitidos por el servidor del editor, porque los derechos de autor en cuestión se refieren a los artículos escritos por los periodistas y los autores científicos y no sobre los otros componentes de la página extraída del sitio de Internet.

Importa poco igualmente que el artículo sea enviado por Google en lenguaje HTML y en seguida transformado por el computador del internauta en página ilegible, porque los servidores de Google copian las páginas en lenguaje HTML, tanto que estas son emitidas en ese lenguaje por el servidor del editor. Es entonces una copia del mismo artículo, escrito en la misma forma que es transmitido por Google al internauta.

Google sostiene sin embargo, que no es quien reproduce y comunica la obra en el sentido de la LDA, pero el internauta cuando hace clic sobre el link <<en caché>> y la descarga en su computador. Google no hace nada más que poner a disposición del internauta una <<instalación>> que lo permitirá hacer una copia.

información alojada en dicha memoria caché, toda vez que la información en ella depositada perdura por incluso años posteriores a su publicación y a pesar de que esta incluso no repose directamente en la página del web master, de manera que dentro de las bases de datos de Google, se puede encontrar información obsoleta o antigua guardada que omite el requisito de transitoriedad.²⁴⁴

22. No se discute que los artículos de los periódicos están beneficiados de la protección establecida por la LDA a título de obra literaria o artística.

En los términos del artículo 1 de la LDA, el autor es el único que tiene derecho de reproducir su obra o de autorizar la reproducción, de cualquier manera o sobre cualquier forma que sea, directa o indirectamente, provisional o permanente, en todo o en parte. Al igual, solo el autor tiene derecho de comunicarla al público por cualquier medio, este comprende la puesta en disposición del público de manera que cada uno pueda tener acceso en el lugar y momento que elija individualmente.

De esta disposición se deduce que el registro por Google en sus propios servidores de una página publicada por un editor constituye un acto material de reproducción. De otra parte, el hecho de que Google permita a los internautas conocer esta copia –que no se debe confundir con la página original- haciendo clic sobre el link <<en caché>> constituye una comunicación al público. En la era digitak, hay reproducción desde el momento donde hay fijación, lo que hace la descarga justificable del hecho de reproducción. (F: De Visseher & Michaux, op. Cit, pág 71, n°88).

Sin la intervención necesaria de Google el internauta no podría tener acceso a esta página porque, al ser una copia de la página tal y como existía al momento en el cual el sitio ha sido visitado por los robots, esta no existe más a nivel temporal cuando el internauta dirige su consulta, en algunos casos esta no existe ni siquiera materialmente cuando esta ha sido retirada del sitio de Internet por el editor.

Es entonces erróneo que Google sostenga que es el internauta quien reproduce los artículos <<en caché>> de los periodistas y otros autores científicos publicados por los editores.

23. Contrariamente a esto que sostiene Google, el artículo 8 del tratado de la OMPI sobre derechos de autor de 20 de diciembre de 1996 no precisa que el servicio <<en caché>> propuesto por el no es una comunicación al público. Por el contrario, este es totalmente silencioso sobre este punto.

En el mismo sentido, el servicio que consiste en permitir tomar conocimiento de una página archivada no puede ser asimilada a <<un simple suministro de instalaciones destinadas a permitir o a realizar una comunicación>> se hace referencia a esta en el considerando 27 de la directiva 2001/29/CE del parlamento europeo y del Consejo del 22 de mayo de 2001 sobre la armonización de algunos aspectos de los derechos de autor y derechos conexos o a fines en la sociedad de la información. En efecto, no conviene confundir la búsqueda instantánea de una página siempre publicada en Internet por medio del servicio normal <<Google Web>> donde Google no juega sino un rol de motor de búsqueda y viene de una página antigua, que existía al momento en el cual ha sido visitada por los robots informáticos de Google que ofrecen así un servicio suplementario.

En conclusión, no hay lugar a asimilar el motor de búsqueda de Google a un simple *copy center* (o centro de copiado) poniendo las fotocopias a disposición del estudiante para permitirles copiar las páginas extraídas de libros o de revistas científicas. En efecto, en el caso concreto, no son los internautas que copian, como si lo harían los estudiantes, sino es Google que pone a su disposición una copia que ellos mismos han realizado”.

²⁴⁴ Ibidem. “El caching o memoria caché o antememoria es, en informática, una memoria que registra temporalmente las copias de datos provenientes de otro recurso de datos, a fin de disminuir

Ahora bien, a pesar que Google en la actualidad se presenta como una herramienta útil, en el ámbito jurídico infortunadamente los resultados de la búsqueda no alcanzan a satisfacer las necesidades de los abogados y demás operadores jurídicos, en vista de la complejidad de las bases de datos jurídicas, que requieren además de fórmulas matemáticas, conocimiento especializado y confiabilidad en la información suministrada, asunto que resulta hoy en día difícil, en vista de la existencia de la web social dentro de la cual no existe barrera alguna para aportar, modificar o crear contenidos web.

el tiempo de acceso (en lectura o en escritura) de un material informático (en general, un procesador) a esos datos. La memoria caché es más rápida y más próxima al material informático que solicita el dato. Los datos incluidos en caché pueden ser por ejemplo un programa, un block de imágenes a tratar. (http://fr.wikipedia.org/wiki/M%C3%A9moire_cache).

El servicio <<en caché>> de Google no puede ser asimilado a un simple almacenamiento en caché o caching informático el cual es asunto de la Directiva 2001/19. Google mismo reconoce que el caché que el propone tiene por finalidad <<Consultar una copia a todo momento, y en particular en el caso en el cual la página original (o Internet) sería inaccesible>>.

De otra parte, Google no demuestra que la inclusión en <<caché>> de los artículos en causa y sobre todo la comunicación al público de sus páginas archivadas es necesaria sobre el plan técnico para asegurar una transmisión eficaz de la obra, sobre todo al afectar notablemente el rendimiento y la veracidad del procesamiento de consultas.

[...]

26. Además, una de las condiciones para beneficiarse de la excepción de copia provisional es que aquella sea <<transitoria>>.

Un acta de reproducción no puede ser calificada de <<transitoria>> al sentido de la segunda condición enunciada en el artículo 5 párrafo 1 de la Directiva 2001/29 solo si su duración de vida es limitada a aquello que es necesario para el buen funcionamiento del procedimiento técnico concerniente, entendiendo que ese procedimiento debe ser automatizado de suerte que este supprime ese acto de una manera automática, sin intervención humana, desde sus funciones que para permitir la realización de dicho proceso se archiva. (C.J.U.E., 16 de julio de 2009, c-5/08, Infopaq, punto 64).

Google no ha respondido que la copia <<en caché>> se encuentre accesible mientras que el editor conserva su artículo en su sitio, este puede durar varios días, semanas, meses o años. De otra manera, esta queda accesible gratuitamente, incluso cuando el editor ha condicionado su descarga al pago de una tarifa y no la publica más en su sitio. Además, es demostrado por la ejecución provisoria del juzgamiento realizado que es posible suprimir la función <<en caché>> de ciertos artículos por una intervención humana.

Con esto se deduce que Google no puede reivindicar el beneficio de la excepción de copia provisional cuando la copia <<en caché>> no sea calificada de transitoria". (*El subrayado no es original*).

Por lo expuesto con anterioridad, es necesario referirnos a la subsistencia de los problemas del pasado de la informática jurídica en la actualidad. La recuperación de la información, la organización de la búsqueda y su clasificación, son dificultades que aún persisten. Motivo por el cual es importante analizar dentro de las herramientas de los buscadores, cuáles fueron las repercusiones de los *search engines* en la informática jurídica.

2.11. Problemas que suscitan los buscadores o search engines en la actualidad

Pese a la gran variedad y especialidad de los buscadores, los resultados de las búsquedas en un sin número de oportunidades no se encuentran acordes con las necesidades y requerimientos demandados por los usuarios que buscan información jurídica. En este acápite entraremos a analizar cuáles son los problemas que se presentan en términos de utilidad, pertinencia, imparcialidad, oportunidad, veracidad e integralidad de la información que arrojan los motores de búsqueda.

El primer problema que encontramos dentro de los buscadores, está relacionado con *la autoridad que realiza la labor de discriminar y organizar la información*, en este sentido, podemos afirmar que dicha actividad las más de las veces la realizan técnicos o ingenieros de sistemas, o en motores de búsqueda, que no cuenta con los conocimientos especializados para determinar la ubicación de la información²⁴⁵, por lo cual en múltiples ocasiones podremos encontrar la presencia de ruidos

²⁴⁵ “Por otra parte, mientras que la obra de los enciclopedistas fue construida por filósofos, científicos y gente de letras, hoy la Internet y otros medios interactivos son el fruto del *knowledge managers*, cuya formación es más técnica que vinculada a las ciencias y las artes, y cuyas metodologías imponen sobre sus productos el halo de neutralidad ideológica tradicionalmente atribuido a los algoritmos computacionales”. SIRI, L. (2000). op.cit. p. 18.

informáticos y ubicación errónea de documentos, lo que conduce a desaparecer la información para siempre, por no encontrarse en su lugar idóneo²⁴⁶.

A propósito del denominado ruido informático, sostiene Laura Siri: “Un libro mal colocado en un librero puede desaparecer para siempre.²⁴⁷” En igual sentido que “El estado actual de los buscadores puede compararse a una guía de teléfono actualizada en forma irregular, con tendencia a listar sólo la información más popular, y que tiene la mayoría de las páginas arrancadas”²⁴⁸.

En este sentido se ha mencionado, que además existe una carencia “en relación con la dificultad de validar la información en Internet. Es decir, la construcción de la “autoridad” en un medio donde cualquiera puede ser tanto productor como consumidor”. Podemos observar claramente, como un buscador puede arrojar contenidos de páginas Web que son de dudosa proveniencia o que sencillamente carecen de un contenido veras, serio y oportuno. En otras palabras, los documentos que reposan en Internet “no están sujetos a los rigores de la industria editorial”²⁴⁹ motivo por el cual la más de las veces encontraremos documentos de origen y autores desconocidos.

Este problema se ve más incrementado cuando quien realiza la búsqueda no conoce los aspectos más relevantes para determinar la confiabilidad del sitio Web, que las

²⁴⁶ Ibídem. p. 28. “No sólo hay en la Red gran cantidad de material cuya clasificación desafía la tradición bibliotecóloga o bien cuya identidad es confusa, sino que muchas búsquedas arrojan como resultado una enorme cantidad de material irrelevante, de autoridad dudosa, inexacto o sin nada que ver con la temática buscada”.

²⁴⁷ Ibídem. p. 13.

²⁴⁸ LAWRENCE, S., & GILES, L. (n.d.). *Metrics.com*. <http://www.wwwmetrics.com/>. Citado por SIRI, L. (2000). op.cit. p. 11.

²⁴⁹ SIRI, L. (2000). op.cit. p. 36.

más de las veces “son más que la corrección del formato y el cumplimiento de ciertos requisitos de *software*, *hardware* y herramientas multimedia. También hay que agregar la existencia de reglas conocidas para la inclusión de la información, la comparabilidad con recursos relacionados, la estabilidad de la información y la autoridad del autor o creador”²⁵⁰.

Otro de los problemas que suscitan los buscadores, consiste en la *dificultad que presenta para el usuario encontrar la información relevante dentro de la gran cantidad de páginas Web que contiene Internet*, motivo por el cual al arrojar el buscador millares de páginas Web sobre las *keywords*²⁵¹ introducidas, se genera una dificultad para el usuario consistente en poder discriminar dentro de lo ya discriminado, cual es la información que realmente se adapta al motivo de su búsqueda.

En este mismo sentido, gracias a un “estudio realizado por los especialistas Ste Lawrence y C. Lee Giles (1999) –del Instituto de Investigaciones NEC en Princeton, Nueva Jersey (EE.UU)- sobre los buscadores demuestra que éstos están lejos de cubrir la totalidad de sitios de la Red y que, además, ofrecen una imagen distorsionada de la misma.”²⁵².

De ahí que, podamos ver la ausencia de protección de identidades locales debido a la inexistencia de las políticas culturales, por lo cual como examinaremos en lo sucesivo, la gran mayoría de páginas con mayor número de visitas son

²⁵⁰ Ibídem. p. 37.

²⁵¹ Key Word, se denomina así a aquella palabra clave introducida dentro del motor de búsqueda con el fin de obtener un resultado determinado.

²⁵² SIRI, L. (2000). op.cit. p. 21.

Estadounidenses, de forma que las otras páginas carecen de difusión sobre toda la red, circunstancia que se encuentra íntimamente ligada a la regionalización de las páginas, teniendo en cuenta la configuración de dominio que tiene el computador, así por ejemplo; si el computador es de Francia su dominio será .fr y sus resultados van a estar enfocados a esta región, de la misma manera sucederá con .mx para México.

Por lo anterior, los efectos de la primacía dentro de Internet de los sistemas de búsqueda provenientes de Estados Unidos, son entre otros los siguientes:

- “1. Al naturalizar sus clasificaciones construyen hegemonía,*
- 2. Desplazan ítems socioculturales críticos en los países subdesarrollados,*
- 3. Atentan contra los patrimonios culturales, técnicos y científicos de éstos,*
afectan sus memorias y desorganizan sus sistemas clasificatorios y
culturales”²⁵³.

Otro problema que presentan los contenidos de los buscadores, consiste en la existencia de los derechos de autor que protege la información contenida en Internet. *El material que queda libre para consultar es residual*, especialmente porque la información valiosa queda en manos de empresas privadas que compran los derechos. Frente a este problema, de ninguna manera pretendemos sostener que la protección a los derechos de autor, no sea ajustada a derecho, por el contrario, es necesario otorgar las garantías, con el fin de salvaguardarlos.

En este sentido, es importante resaltar “la necesidad de que las políticas públicas relativas al uso de Internet estén basadas en un análisis realista de los desafíos

²⁵³ Ibidem. p. 40.

planteados por las nuevas tecnologías al corpus jurídico preexistente, y no en un mero (y, a menudo, sólo discursivo) impulso a la generalización del acceso”²⁵⁴.

Otra deficiencia que se evidencia dentro de los buscadores, consiste en *la cuestionable calidad de la información suministrada por los mismos*, hecho palmario que cuenta con el respaldo de los temas anteriormente mencionados, sobre el pago por posicionamiento dentro del PageRank, la existencia del spam, la prevalencia de páginas Web provenientes de Estados Unidos, que generan la regionalización de la información y su parcialización²⁵⁵, las innumerables cantidades de sitios Web que contienen información comercial y la veracidad de la información contenida en las páginas que carecen en varias ocasiones de fuentes bibliográficas formales que respalden el contenido de sus afirmaciones.

En este sentido, mediante el estudio sobre los buscadores que realizaron los especialistas de Ste Lawrence y C. Lee Giles en 1999, en el Instituto de Investigaciones NEC en Princeton, Nueva Jersey (EE.UU), se evidenció que “los sitios comerciales constituían el 83 por ciento de la Internet públicamente accesible y que los buscadores albergaban sólo un 6 por ciento con contenidos educativos o científicos. Este resultado se puede contrastar con un informe del 25 de octubre de 1999, publicado por AltaVista, según el cual sólo el 20 por ciento de todas las

²⁵⁴ Ibidem. p. 41.

²⁵⁵ Ibidem. p. 25. “Este es un proceso antiguo pero en el cual hoy pesa la hegemonía ‘ideológico/clasificatoria’ de Estados Unidos presente en toda búsqueda en Internet y produce, como también sucede en las enciclopedias en CD-ROM, un proceso de desinformación global importante con respecto a los países pobres cuando no una intermediación o intrusión en la información sobre ellos errónea e hiper-simplificada. (...)”

En efecto, la forma en que la información que circula en Internet es clasificada, jerarquizada, puesta en primer plano o bien ocultada por esos sistemas informáticos con pretensión de universalidad que son los buscadores, en realidad naturaliza el vaciamiento de las memorias y culturas de, principalmente, los países menos desarrollados”.

búsquedas llevadas a cabo por su servicio están relacionadas con productos comerciales”²⁵⁶.

Ahora bien, con el fin de mencionar otro problema presentado por los buscadores, presentaremos a continuación, las tres dimensiones de dificultades que tiene el intercambio de información en la web:

“La expansión de la información; los distintos aspectos de la calidad de esta y los problemas relativos a su vigencia”²⁵⁷.

Examinado lo anterior, el inconveniente que presentan los buscadores en Internet, consiste en *la gran diferencia con las técnicas tradicionales de bibliotecología para clasificar y recuperar la información*. “Internet –o la web- no fue diseñada para dar soporte a una publicación ni a una recuperación de la información, en forma organizada, como si lo fueron las bibliotecas:

*“Este almacén de información contiene no solamente libros y papers sino también datos científicos en crudo, menús, agendas de reuniones, anuncios publicitarios, grabaciones de audio y video y transcripciones de conversaciones interactivas. Aunque la Red no sea una biblioteca, su continuo crecimiento hace necesario organizar, algo parecido a los tradicionales servicios bibliotecológicos para acceder y preservar la información.”*²⁵⁸.

²⁵⁶ Ibidem. p. 22.

²⁵⁷ Ibidem. p. 44.

²⁵⁸ Ibidem. p. 45.

En consecuencia, la información contenida en internet, debería contar con las técnicas tradicionales de la bibliotecología, por lo cual sería importante que los documentos cuenten con la totalidad de su información bibliográfica, esto es; la fecha de creación, el autor del documento y en general todas las fuentes formales, de manera que se pueda acceder a la información de forma completa y eficaz.

Los buscadores, al igual que la labor de bibliotecología, deberán cumplir con las cuatro fases: “Recopilación, selección y adquisición de documentos, análisis documental (dentro del cual se debe incluir el análisis de contenido y el establecimiento de las características físicas del documento y de su soporte), búsqueda, recuperación o investigación documental y difusión.”²⁵⁹.

Por lo anterior, la labor de selección y control de la información en Internet, reviste total importancia, toda vez que la cantidad de información que reposa allí es ostensible, de forma que se torna absolutamente necesario e imprescindible organizarla con el fin de no optimizar su contenido y otorgar una mayor seriedad al mismo²⁶⁰.

En este sentido, se ha mencionado que el hecho de haber remplazado a los humanos y automatizar la organización y clasificación de la información, tiene por causa principalmente la disminución de los costos en los cuales tendrían que incurrir, de

²⁵⁹ Ibidem. p. 47.

²⁶⁰ Ibidem. “Existen tres tipos de abordajes para el control de la información en Internet: primero, la indización o indexación automática de los motores de búsqueda, segundo, la indización manual que proveen los *clearinghouses* de recursos en una o varias áreas del conocimiento y, por último, la catalogación de los materiales de Internet tal como lo realiza OCLC (*Online Computer Library Center*) en Estados Unidos.

tratarse de personas con conocimientos especializados, motivo por el cual se llega en múltiples ocasiones a parcializar la información, incurriendo en errores que desembocan en resultados de búsqueda equívocos o incompletos.

A propósito de esto, sostiene Laura Siri lo siguiente:

“La búsqueda, con frecuencia, produce referencias a sitios web irrelevantes, mientras que son dejados de lado materiales que quien hace la indagación hubiese considerado importantes. Puede decirse que, “así como una biblioteca es tan buena como el catálogo de sus libros, la World Wide Web es tan útil como los search engines que le dan servicio”²⁶¹.

Seguidamente, encontramos que otro problema que presentan los buscadores consiste en que *los buscadores arrojan resultados indiscriminados de todo lo existente en la Web*, por lo cual, se requieren bancos de datos especializados que contengan la información seleccionada. Lo anterior, con el fin de prevenir que se desdibuje y se desoriente la búsqueda introducida por medio de los *keywords*.

Como mencionábamos dentro del concepto de *search engine*, existe gran variedad de documentos que integran el contenido de la Web, así pues, podremos encontrar audio, video, documentos en texto, juegos, programas, entre otros. De ahí, que surja un problema relacionado con *la dificultad de manejar con criterios bibliotecólogos los contenidos de la Web, mediante el uso de los buscadores o search engines*.

²⁶¹ Ibidem. p. 48.

Lo anterior, a causa de la dificultad para establecer la identidad, características y contenido de los objetos de la Web. Igualmente, el problema del manejo de contenido de la información se incrementa debido a que pese a la existencia de millares de páginas Web, la gran mayoría de ellas carecen de amplios contenidos sobre un determinado tema, por lo cual, son pocas las ocasiones dentro de las cuales al introducir una *Keyword* el usuario va a encontrar dentro de un *search engine* un libro o publicación académica de contenido técnico específico sobre el tema objeto de búsqueda.

En igual sentido, se presentan vicisitudes relacionadas con las *keywords* al momento de la introducción de la búsqueda, por cuanto la gran mayoría de los usuarios de buscadores no cuentan con conocimientos bibliotecólogos indispensables, para poder dirigir sus búsquedas y encaminarlas por medio de un lenguaje documental.

En este sentido, incluso quienes conocen el uso de los operadores booleanos, estudiados en los antecedentes de los *search engine*, se encuentran con un mal funcionamiento de los mismos dentro de los motores de búsqueda, toda vez que al emprender la búsqueda incluyendo operadores para su restricción no se arrojan los resultados esperados.

Verbigracia: En Google, al incluir dentro de los resultados de búsqueda “LICENCIA DE MATERNIDAD” “DESPIDO SIN JUSTA CAUSA” “FUERO MATERNO” arroja 6.250 resultados, no obstante, si se incluye dentro del campo de búsqueda una palabra clave para restringir la búsqueda como por ejemplo: “LICENCIA DE MATERNIDAD” “DESPIDO SIN JUSTA CAUSA” “FUERO MATERNO”

“SEGURIDAD SOCIAL” arroja exactamente la misma cantidad de resultados, esto es 6.250.²⁶²

Este ejemplo sería incompleto, si dejamos de lado otro hecho que denota la inestabilidad y variabilidad de los resultados de la búsqueda en los motores como Google, este consiste en que al emprender la búsqueda al día siguiente, los resultados aumentan considerablemente pasando de 6250 resultados a 7540 para el caso de la búsqueda “LICENCIA DE MATERNIDAD” “DESPIDO SIN JUSTA CAUSA” “FUERO MATERNO” y para la búsqueda con restricciones “LICENCIA DE MATERNIDAD” “DESPIDO SIN JUSTA CAUSA” “FUERO MATERNO” “SEGURIDAD SOCIAL” los resultados fueron 7530.²⁶³

Otro de los problemas, que en nuestro concepto es catalogado como uno de los más delicados consiste en que “Los buscadores arrojan resultados sobre las búsquedas realizadas en sus propios índices y no sobre la web directamente. Esto último sería imposible debido al volumen de información y al tiempo requerido para realizar la búsqueda. El rastreo de páginas web que efectúan los robots se realiza con periodicidad variable y, por economía, tiende a no ser muy frecuente. Por ello, siempre existe una diferencia entre lo que pueden recuperar los buscadores y lo que realmente se encuentra en la web.”²⁶⁴

²⁶² Consultas realizadas el día 14 de febrero de 2012 en el buscador de Google, cuya dirección web utilizada fue: www.google.com.co.

²⁶³ Consultas realizadas el día 15 de febrero de 2012 en el buscador de Google, cuya dirección web utilizada fue: www.google.com.co.

²⁶⁴ MORENO JIMÉNEZ, P. M. (2005). *Estrategias y mecanismos de búsqueda en la Web Invisible*. p. 1-5. Bogotá: Universidad Militar Nueva Granada. Consultada en diciembre 13, 2011. <http://www.umng.edu.co/www/resources/web%20invisible.pdf>

Esto trae como consecuencia, que existan páginas en la World Wide Web que no se puedan recuperar a través de los buscadores generales, a este conjunto de páginas se les ha denominado web invisible, web profunda o *deep web* y contrario a lo que se podría pensar, la cantidad de páginas que la conforman asciende de un tercio a la mitad de los documentos con que cuenta en su totalidad Internet²⁶⁵, de otra parte y no menos importante, una característica que las identifica, consiste en que gran parte de sus contenidos son de alta importancia y valor para los usuarios, toda vez que en ocasiones sus contenidos gozan de la mayor seriedad.

Podemos definir la Web invisible, como el conjunto de información que reposa en la world wide web, que no se encuentra identificada o detectada por el web crawler, spider o araña de los diferentes motores de búsqueda, de manera que esta permanece oculta para el usuario que emprende una búsqueda, por lo cual los resultados se tornan incompletos, esta información que aparece como inexistente, puede formar parte de todo o parte de un archivo contenido en una página web.

Otra definición que se ha realizado sobre la web invisible, establece que es “información que no puede recuperarse con los mecanismos de búsqueda comunes. Estos mecanismos tratan de abarcar toda la web, pero se calcula que los mayores

²⁶⁵ *Ibidem*. “A pesar de su pretendida exhaustividad, se calcula que los mayores motores de búsqueda indizan sólo entre un tercio y la mitad de los documentos disponibles al público en la Red (Turner, 2003).

Un estudio muy citado de Bright Planet ha estimado que la información contenida en la web invisible es aproximadamente 550 veces mayor que la de la web superficial y crece a mucha mayor velocidad (Bergman, 2000). Un estudio posterior de Cyveillance calcula que el tamaño de la web profunda es 275 veces mayor que el de la web visible (Murray, 2000, citado en Ouf, 2001).

Estimaciones posteriores, en cambio, señalan que el tamaño de la web invisible es sólo entre 2 y 50 veces mayor que el de la web visible (Sherman y Price, 2001b). Las diferencias en las cifras se deben a las diferentes metodologías utilizadas por los autores. En cualquier caso, el valor de la información contenida en la llamada web profunda justifica su estudio y el de sus formas de acceso”.

motores de búsqueda alcanzan a indizar sólo entre un tercio y la mitad de los documentos disponibles. La web invisible no sólo es de mayor tamaño que la web visible o superficial sino que crece a mayor velocidad. Así mismo, mucha información disponible en la web profunda, como la que se encuentra en bases de datos, tiene un alto valor potencial para el usuario”²⁶⁶.

En igual sentido la web invisible, ha sido definida como aquellas “páginas de archivos de texto, u otra información de alta calidad de autoridad disponible a través de la World Wide Web que, los motores de búsqueda de propósito general no pueden, debido a limitaciones técnicas, o no quiere escoger debido a una elección deliberada, adicionar a sus índices de páginas web. A veces también se refiere a la “web profunda” o “materia oscura”²⁶⁷.

Las causas del surgimiento de la web invisible son múltiples, dentro de ellas podemos citar como las más relevantes las siguientes: Que la página web no haya sido indizada, que el sitio web forme parte de la web privada y excluya el acceso a contenidos por conducto de *passwords* o que incluya un campo *noindex* para evitar que el buscador pueda indicar el cuerpo de la página, que la página forme parte de la denominada web propietaria, dentro de la cual es indispensable registrarse con el fin de tener acceso a los contenidos, bien sea de forma onerosa o gratuita y finalmente encontramos como última causa del surgimiento de la denominada web invisible, a la propiamente dicha web invisible, que se compone de páginas que no pueden ser indizadas ya no debido a las especificaciones de su creación, sino a las limitaciones

²⁶⁶ Ibídem. “El término “web invisible” fue utilizado por primera vez por la Dra. Jill Ellsworth para denominar la información que resultaba “invisible” para las maquinarias de búsqueda”.

²⁶⁷ SHERMAN, C., & PRICE, G. (2007). *The invisible web. Uncovering information sources search engines can't see*. New Jersey: Information Today. p. 57.

técnicas de los web crawlers o spiders y en general a las condiciones de los buscadores.

Una de las causas determinantes para la existencia de la web invisible y en especial, del considerable volumen de información que reposa en la web con difícil acceso para los usuarios, consiste en la individualidad que pregonan cada uno de los buscadores, toda vez que cada uno de ellos cubre un segmento del espacio web y estos segmentos no se unifican, dado que como resulta lógico, el giro ordinario del negocio se centra en el suministro de la mayor cantidad de información, por lo cual todos son competencia en el mercado.

En este sentido, además de los buscadores, existen los denominados metabuscadores o *metasearch engines*, que pretenden al igual, abarcar la máxima cantidad de información contenida en la world wide web y eliminan duplicaciones de resultados de los buscadores, de forma que unifican los índices de cada buscador, verbigracia de ellos es: Dogpile, Mamma, Metacrawler, InfoSpace. Estos buscadores a pesar que presentan un resumen de los resultados, cuentan con un inconveniente, toda vez que al no presentar opciones de búsqueda avanzada se afecta la precisión de los resultados.

De igual forma existen los denominados directorios, que a diferencia de los buscadores que utilizan web crawlers o spiders automáticamente o de los metabuscadores, emprenden la labor de recopilación de información de forma manual a partir de un registro que realiza cada autor dentro del buscador, de manera

que esta información se clasifica y jerarquiza para así lograr su organización. Como ejemplo de estos directorios, bien podemos citar a LookSmart y Yahoo!.

Los directorios, que trabajan de la mano con los editores, realizan una labor que si bien cuenta con costos mayores, ofrecen una mejor calidad de la información, no obstante en algunos casos los editores de estos directorios son recursos voluntarios, por lo cual, cuando el modelo se estructura de esta forma, se le denomina modelo abierto o *open model directories*, que disminuye los costos, pero que proporcionalmente, también disminuye la calidad de la recopilación.

De igual forma, otro problema de los directorios, consiste en que estos al estar limitados por la manualidad de sus operaciones, por razones de tiempo, de habilidad del recurso editor que maneja el contenido y a causa de conocimientos limitados sobre una ciencia específica, por ejemplo: El derecho, que ocasiona que la información contenida sobre el tema en cuestión resulte poco representada o excesivamente saturada de información impertinente para el tema objeto de búsqueda.

Otras estrategias, que se emplean para emprender búsquedas en la web invisible de información valiosa, consisten en la combinación de sistemas. Los buscadores, incluyen sistemas de navegación en forma de directorios con el fin de orientar la búsqueda dentro de la base de datos preexistente. Por su parte, los directorios, complementan su información con los de la búsqueda en un buscador específico, de manera que la combinación de herramientas otorga la posibilidad de reducir la existencia de la web invisible. Sin embargo, cuando se emprenden “las búsquedas en

las categorías de un directorio elaborado para un buscador, o las búsquedas por palabras en un directorio, se presentan ciertas limitaciones²⁶⁸.

Una vez examinado lo anterior, es claro que tanto el mecanismo de los directorios, como el de los buscadores, ofrecen ventajas a la hora de emprender una búsqueda en la web, no obstante para lograr abarcar la mayor cantidad de información que reposa en la web, es recomendable combinar las diferentes opciones, toda vez que como se puede dilucidar de lo anteriormente mencionado, los directorios ofrecen una opción de navegación mucho más recomendable, mientras que a la hora de examinar una key Word o palabra clave determinada, es recomendable emprender la búsqueda con el uso de un search engine o buscador.

Finalmente, en cuanto a la crítica de la web invisible, es de anotar que varios son los planteamientos sobre las posibles soluciones para atenuar su existencia, una de ellas, quizás la más acogida, consiste en la propuesta de crear un registro de cada uno de los sitios web que reposan en Internet, de manera que fuese una obligación para el propietario, realizar esta inscripción y así poner a disposición de todos los usuarios el contenido del mismo, de esta forma no existirían URLs desconectadas o invisibles para los buscadores; con la salvedad que la información aportada de la página web deberá respetar los límites de privacidad, que en virtud de la autonomía privada, se hubieren establecido.

Otro cuestionamiento que se le ha realizado a los documentos contenidos dentro de la Web, consiste en que *existe una gran cantidad de documentos de carácter temporal*, motivo por el cual su búsqueda se torna más engorrosa y en ocasiones

²⁶⁸ MORENO JIMÉNEZ, P. M. (2005). Op. cit.

imposible. Por ello, podemos encontrar páginas creadas para un evento, situación o momento específico, por lo cual la información contenida en la Web a diferencia de las bibliotecas convencionales, tiene como una de sus características principales, la mutabilidad o inestabilidad de la información en ella contenida.

Al respecto se ha mencionado que “Una muestra al azar de páginas sobre 200 sitios a lo largo de tres meses en 1995, hecha por dos universidades de Estados Unidos como parte de una investigación sobre un nuevo sistema de indexación, indicó que el tiempo promedio en que una página permanecía sin cambios en la Web era de solo 75 días. Un porcentaje sustancial cambiaba cada diez o menos días. A veces las páginas desaparecían completamente pero, más a menudo, la información que contenían simplemente era actualizada, o bien la página se movía hacia una dirección diferente²⁶⁹.”

No obstante, en la actualidad existen buscadores que contienen opciones de búsqueda avanzada dentro de las cuales se puede recuperar información antigua que se encuentra en la denominada memoria cache²⁷⁰. Sin embargo, no en todos los casos

²⁶⁹ SIRI, L. (2000). op.cit. p. 57.

²⁷⁰ “**Cache-control: public** – indica que la versión cacheada puede ser guardada por proxies y otros servidores intermedios para que todo el mundo tenga acceso a ella”.

Una web puede ofrecer tres tipos de contenidos en función de su cacheabilidad:

- 1. Contenidos estáticos:** Son aquellos contenidos que ya existen físicamente en el servidor, tienen un tamaño fijo y una fecha determinada. Una página html, una imagen o un documento pdf son ejemplos de contenidos estáticos. Este tipo de contenidos son firmes candidatos a ser cacheados.
- 2. Contenidos dinámicos:** Son contenidos que se crean en el momento en el que se solicitan accediendo, por ejemplo, a una base de datos. Este tipo de páginas son independientes del perfil del usuario que las visita y por tanto son potencialmente cacheables.
- 3. Contenidos personalizados:** Este tipo de contenidos se crean a medida en base a las características de la petición y pueden ser diferentes en función de *quién*, *cuándo*, o *cómo* es solicitada. Por este motivo este tipo de contenidos no son buenos candidatos a ser cacheados, aunque como veremos más adelante hay técnicas que permiten cachearlos de manera parcial o

se podrá recurrir a esta herramienta obteniendo resultados satisfactorios, dado que en algunas ocasiones son sus propietarios quienes toman la decisión de suprimir tal o cual información, además de tener en cuenta la naturaleza específica del tipo de documento que se pretenda cachear.

En igual sentido, es pertinente recordar lo mencionado a propósito de los mecanismos de recuperación de información de Google, que en lo relacionado con la memoria caché, logra conservar contenidos incluso desuetos de años atrás, por lo cual al emprender consultas la información no resulta oportuna y altera los resultados e impresión que sobre determinado asunto el usuario pudiese tener.

De otro lado, hay quienes afirman que “los actuales “árboles del conocimiento.” Propuestos por algoritmos de búsqueda promocionados por sus fabricantes como instrumentos con la potencialidad técnica de incluir todo tipo de fuentes por igual- en realidad tienen algunas ramas “podadas””²⁷¹. Esto encuentra justificación, toda vez que los servicios de búsqueda actuales no son realmente buscadores sino directorios que en sus primeras páginas ubican las páginas más visitadas, las que han logrado un posicionamiento en la web a causa de su alta indexación o aquellas que han cancelado determinada suma para posicionarse dentro de los resultados.

total. El área privada de la web de un banco, o un carrito de la compra son dos ejemplos de páginas personalizadas.

Es muy importante que entendamos qué contenidos pueden y deben cachearse y cuáles no. Si no gestionamos eficientemente el cacheo corremos el riesgo de servir a nuestros usuarios un contenido que no está actualizado o por otro lado saturar nuestro servidor con peticiones innecesarias”. GOROSTIZA, I. (2009, Mayo 25). *Tutorial Cache Web: cómo gestionar el cacheo de nuestros contenidos*. [hellogoogle.com: http://www.hellogoogle.com/tutorial-cache-web/](http://www.hellogoogle.com/tutorial-cache-web/)

²⁷¹ SIRI, L. (2000). op.cit. p. 16.

En este estado de cosas, podemos concluir, en lo que atañe a la información jurídica, que los buscadores de acceso abierto como Google, Bing y Yahoo, etc. no satisfacen las necesidades de usuarios de información jurídica que deseen encontrar información confiable, organizada, fidedigna y veraz, toda vez que como ya lo mencionamos, son innumerables las falencias que presentan.

Ahora bien, es de reconocer que existen buscadores especializados que contienen bases de datos jurídicas de acceso libre o restringido por concepto de costo de inscripción y cuentan con altos y confiables contenidos de información jurídica, tales como Legifrance, CREDOC (Centre de Documentation Automatique du Droit), Editorial la Ley, entre muchos otros bancos de datos que existen a nivel mundial, no obstante estas bases de datos limitan su tarea a la recopilación de la información, que la más de las veces al ser examinada por el usuario se torna difícil de consultar, a causa del trabajo engorroso que se debe realizar para encontrar el descriptor adecuado con el cual se incluyó el documento objeto de búsqueda. Estas bases de datos, en el mayor de los casos, obedecen al concepto de la Web 1.0 y se crean en formato HTML.

2.12. Buscadores temáticos

Los buscadores temáticos, son aquellos que proporcionan información que concierne a determinada materia, sin que se discrimine la información asociada con el tema objeto de la búsqueda. Estos resultados de la búsqueda, se arrojan organizados de forma jerárquica, figurando en orden de prioridad los generales y luego los

específicos, además de clasificar por categorías sus contenidos y en algunos casos realizar comentarios descriptivos sobre los mismos.

En el mismo sentido, se han definido como aquellos que “proporcionan información directamente relacionada con una determinada materia, a diferencia de los buscadores generales, en los que a la introducción de un término de búsqueda el buscador responde con todas las entradas que tenga almacenadas en su base de datos, sin discriminar las materias a que están asociadas. El filtrado de páginas que se realiza en un buscador especializado supone un considerable ahorro de tiempo para el usuario”²⁷².

En lo relacionado con el objetivo de los buscadores temáticos, este consiste principalmente en encontrar los documentos relacionados de forma directa con el tema a cerca del cual se emprende la búsqueda. Esta clase de buscadores se encuentran compuestos por dos partes a saber: La primera, una estructura asimilada al tesoro que proporciona la lista organizada con una estructura jerárquica que permite una consulta más fácil para el usuario de la misma y de otra parte los buscadores temáticos cuentan con una base de datos propia que se construye por el conjunto de URLs remitidos.

De otra parte, el funcionamiento de los buscadores temáticos, se encuentran conformados además de la estructura jerárquica anteriormente mencionada, por la división de temas que van desde lo más genérico a lo más específico, de forma que

²⁷² TEMPRANO SÁNCHEZ, A. (2008). op. cit. p. 53.

existe continuidad desde el inicio del tema hasta asuntos puntuales donde probablemente se podrá encontrar la información solicitada.

Dicho lo anterior, cabe resaltar que la información se encuentra estructurada como un directorio y se ofrece al usuario dos opciones de búsqueda: La primera, que consiste en la utilización de un buscador automático de directorios y contenidos y la segunda, el uso de las listas prolongadas, en atención al orden y ubicación general o específica.

Al igual que los buscadores generales o automáticos, los buscadores temáticos, cuentan con varias deficiencias a saber:

Entre ellas, cabe mencionar que la búsqueda en varias ocasiones se torna dispendiosa toda vez que exige cumplir con varios pasos antes de localizar la información solicitada.

En segundo lugar, encontramos que los buscadores temáticos, al contar con listas prolongadas que van de lo general a lo específico, en ocasiones no se encuentran actualizadas, por lo cual las búsquedas no podrán arrojar la información oportuna.

Al existir unas listas con gran tamaño, puede ocurrir que en el entre tanto de la búsqueda, el usuario se disperse hacia otro tema que atraiga su atención.

Al crear un listado que señala unos temas catalogados como generales, resulta obligatorio encuadrar la totalidad de los temas en alguno de estos títulos generales,

por lo cual, existen temas que resultan de difícil categorización, lo que trae consigo que el material pueda quedar en una ubicación errónea.

2.13. Buscadores especializados

Una vez estudiados los buscadores genéricos que pretenden localizar información sobre todos los temas contenidos en la World Wide Web, es menester mencionar que de forma paralela a estos, se encuentran los buscadores que suministran información sobre temas particulares. Los buscadores especializados, son un conjunto de documentos web encuadrados dentro de una base de datos, que proporcionan una herramienta útil para el usuario especializado en un área puntual del conocimiento. Estas se crean con el fin de administrar conjuntos y tipos de información. Son búsquedas enfocadas sobre un área particular y se crean gracias a un interés diverso que no se satisface en la larga lista de bases de datos y buscadores generales existentes.

Otros autores, han definido a los buscadores especializados como “aquellos que en su propia base de datos solo contienen direcciones que han sido más o menos filtradas y que se corresponden en mayor y menor medida con el sentido general de nuestras necesidades de información.”²⁷³.

Estos buscadores, surgen como respuesta a la necesidad de crear búsquedas especializadas sobre un tema en particular y tienen gran utilidad a la hora de hablar de contextos educativos y profesionales, toda vez que se circunscriben a un área

²⁷³ MUÑOZ LÓPEZ, M. (2010). op. cit. p. 1994.

determinada, de manera que al arrojar la información se logran mejores resultados que en los buscadores genéricos, esto sin pretermir que el usuario no tendrá que invertir gran cantidad de tiempo en discriminar la información dentro de los resultados de la búsqueda.

Por lo anterior, los buscadores generales son sustituidos por los buscadores especializados siempre que se pretenda emprender una búsqueda sobre un dominio o área concreta, toda vez que estas herramientas limitan el ámbito de la búsqueda y proporcionan a los usuarios contenidos de alta calidad sobre las páginas que se encuentran en la web.

Una de las ventajas, que podemos resaltar en este tema, consiste en que a diferencia de los buscadores generales que proporcionan información útil siempre que se trate de búsquedas sencillas, los buscadores especializados como su mismo nombre lo dice, otorgan resultados específicos y concretos, sobre temas que revisten y requieren unas características especiales.

Como características de los buscadores especializados, podemos mencionar, entre otras las siguientes:

- Utilizan un vocabulario controlado, rigurosamente formal Manejo de la información de forma discriminada.
- Contenido de los documentos enfocado al tema objeto de la base de datos
- Restringen los resultados de la búsqueda al área del conocimiento que se sistematiza.

Los buscadores especializados cumplen una importante labor, toda vez que al emprender la búsqueda se garantiza que el tema objeto de la misma, se limite al contenido del buscador especializado en la materia. Al introducir la palabra clave el usuario encontrará dentro del resultado de la búsqueda únicamente información correspondiente al tema objeto de ella. En este punto es importante mencionar que todo esto se logra gracias a una ardua tarea que implica la inversión de más tiempo y mucho más trabajo que los métodos de los buscadores genéricos y temáticos que resultan más onerosos y complejos en su elaboración y no satisfacen las necesidades del usuario jurídico.

Ahora bien, en los buscadores especializados, los resultados a diferencia de los motores de búsqueda generales como Google, no se encuentran posicionadas por criterios de consulta, es decir que si un usuario ingresa a una página web y la consulta por una gran cantidad de tiempo, esto no va a determinar la importancia o no de la página, de igual forma la cantidad de citas que se realicen en otras páginas web no cobrarán importancia a la hora de arrojar los resultados de las búsquedas, toda vez que los criterios de relevancia corresponden a la especificidad del tema.

En igual sentido, una de las ventajas de los buscadores especializados, consiste en que en gran cantidad de oportunidades, estos buscadores, logran la indexación de páginas web de su especialización que no se localizan en buscadores generales y que incluso en ocasiones se encuentran ubicadas en la denominada *web invisible*.

Si se tiene en cuenta lo anterior, en términos de confiabilidad, podemos sostener que el uso de los buscadores genéricos trae consigo defectos inherentes, tales como la deficiencia en los criterios con los cuales se establece la relevancia de una página web, por cuanto existe la posibilidad de dejar transcurrir el tiempo con varias ventanas abiertas sin encontrarse consultándola, hecho este que puede llegar a otorgarle un mejor posicionamiento a dicha página sin conocer realmente si cuenta o no con la relevancia para obtener una mejor ubicación dentro del buscador.

2.14. Antecedentes de los sistemas de recuperación de información jurídica

Antes de iniciar el estudio de los antecedentes de los sistemas de recuperación de información jurídica en el presente, es pertinente remitirlos a la primera parte de esta tesis doctoral, con el fin único de conocer sus orígenes y el pasado de los mismos antes del surgimiento de Internet, toda vez que como veremos a continuación, se hará referencia a lo que sucedió con la aparición de la world wide web e Internet.

Ahora bien, la historia de los sistemas de recuperación jurídica que surge a partir de la creación de Internet y la World Wide Web, se remonta a la creación de CELEX (Communitatis Europae Lex), que surgió para lograr mayor efectividad con la creación de mayor cantidad de campos de búsqueda.

Con el surgimiento, de sistemas que incluían ampliación en los campos de búsqueda, hubo una transición a la creación de formatos que incluso empezaron a contar con procesadores de texto y otorgaron la posibilidad de crear lenguajes de marcación de la información jurídica para fines documentales, estos lenguajes de marcas, permiten

la codificación de los documentos, e incluyen etiquetas o como su nombre lo dice, marcas, que contienen la información sobre la estructura del texto.

Como se mencionó en el acápite sobre la web 1.0, al hacer referencia a HTML "HyperText Markup Language o lenguaje de marcado de hipertexto", este es el primer formato con el cual contó la web y el más extendido, no obstante como se pudo examinar con anterioridad, existen otros. Es de aclarar que los lenguajes de marcado no cuentan con funciones aritméticas, toda vez que no son lenguajes de programación.

“A finales de los años sesenta del siglo anterior, tres investigadores contratados por IBM, Charles Goldfarb, Ed Mosher y Ray Lorie, recibieron el encargo de diseñar un sistema de edición, almacenamiento, búsqueda y gestión de documentos legales construyendo un sistema de formateo estructural al que, en un principio, se denominó GML. En 1986 se convertiría en un estándar, SGML —Standard Generalized Markup Language—, metalenguaje de etiquetado de texto convertido en norma ISO 8879.

Marta Blaquier (1996) subraya cómo los proyectos relevantes de marcado aplicados a la documentación legal no se presentaron hasta la década de los noventa, de la mano de editores privados como Kluwer o Thomson Legal Publishing y con importantes iniciativas acometidas por organismos públicos: la Government Printing Office, de Estados Unidos; la Her

Majesty's Stationery Office, editora oficial del Reino Unido, y la Office for Official Publications (Publications Office) de las Comunidades Europeas.

Igualmente reseñables resultan proyectos como el CETL (Center for Electronic Text in the Law), de la Universidad de Cincinnati, así como los proyectos europeos de Finlandia y Suecia; en este último país resulta obligado hacer una referencia al Corpus Legis Project, nacido en la Universidad de Estocolmo con el fin de elaborar recursos de textos electrónicos legales para la realización de estudios jurídico-lingüísticos”²⁷⁴.

Si se tiene de presente, que los buscadores generales pretermiten las peculiaridades del manejo del documento jurídico, en virtud de lo expuesto al hablar de su problemática, resulta claro e incuestionable reconocer que si bien las herramientas tecnológicas de la web son innumerables y provechosas, en particular con lo relacionado con el lenguaje de marca, estas aplicaciones deben tener de presente las particularidades del manejo documental de la información jurídica.

A partir de lo anterior, se dilucidaron bases de datos jurídicas que iniciaron a implementar el lenguaje de marcado, dentro de ellas podemos citar “Norm@civil — proyecto del área de Derecho Civil de la Universidad de Gerona que incluye legislación y jurisprudencia—, Noticias Jurídicas—portal de la editorial Bosch que contiene normativa concordada, artículos doctrinales o guía judicial—, el portal IUStel, algunos boletines oficiales autonómicos”²⁷⁵.

²⁷⁴ ALVITE DÍEZ, M. L. (2009). op.cit. p. 36.

²⁷⁵ *Ibíd.*

De otra parte, en atención al surgimiento de la web 2.0 que llegó acompañada de nuevos formatos como el XML, se propende por la creación de un sistema integrado que conjugue lo que ya contenía el formato HTML, pero con mayor velocidad y que otorgue herramientas semánticas a la información jurídica.

Con el surgimiento del formato XML, se crea la posibilidad de definir los tipos de documentos y la creación de etiquetas para su codificación y posterior procesamiento y exploración automática, que para los efectos de la informática jurídica permite la recuperación de la información.

Con el formato XML, se proporcionó a la informática jurídica la potencialidad requerida para el tratamiento de la información. “Son abundantes los ejemplos de aplicación de esta tecnología. Conviene recordar proyectos emprendidos en Canadá o Estados Unidos, como el Electronic Court Filing Project. Este grupo de trabajo se unió con otro de la Universidad de Utah con el fin de debatir y desarrollar DTD de XML para la edición e intercambio de documentos jurídicos, y con Gabe Wachob, del portal jurídico FindLaw, un grupo considerado el promotor de la organización internacional Legal XML”²⁷⁶.

En el año de 1998, en Estados Unidos, como respuesta a los avances en la implementación del lenguaje XML y a la creencia que la aplicación generalizada del documento y la información, depende del intercambio de la tecnología en el desarrollo y la adopción de medidas unificadas, surgió LegalXML, como una organización, que reúne a expertos abogados, administradores de tribunales,

²⁷⁶ Ibídem. p. 38.

desarrolladores, agencias gubernamentales, miembros de la academia, empresas, proveedores de aplicaciones y técnicos informáticos, con la única finalidad de unificar, crear estándares para el intercambio electrónico de datos jurídicos en formato XML y de tecnología en general²⁷⁷.

En sus inicios, LegalXML, se centró en la presentación electrónica de documentos de la Corte, con la finalidad de resolver las múltiples preocupaciones del manejo de la información jurídica para los abogados y los administradores de los registros de casos judiciales.

En el año 2002, con el fin de aprovechar el apoyo otorgado por OASIS, principal organización del mundo en XML, en vista, que muchos de los trabajos realizados en LegalXML, se encontraban relacionados con las múltiples preocupaciones de las organizaciones y asociaciones nacionales a parte de OASIS, se presenta a consideración, con el objeto de evaluar y eventualmente aprobar los avances sobre la información jurídica en XML y establecerlos como estándares comerciales y técnicos.

En conjunto con especificaciones dadas por Electronic Court Filing Technical Committee, se renviaron para su revisión y se adoptaron por parte de la National Association for Court Management (Asociación Nacional para la Administración de la Corte NACM) y por el Joint Technology Committee of the Conference of State

²⁷⁷ *Oasis LegalXML*. (2005, Mayo). <https://www.oasis-open.org/committees/download.php/12785/OASIS%20LegalXML%20Brochure.pdf>

Court Administrators (Comité de Tecnología Conjunta de la Conferencia de Administradores de Tribunales del Estado COSCA).

En lo relacionado con la misión de LegalXML, encontramos que en vista de la estandarización que ofreció, en la actualidad continúa con la definición de las normas técnicas para la estructuración de los documentos legales y la información por conducto de las tecnologías XML y relacionadas. Dentro de ellas contribuye en temas como el e-government, firma digital y comercio electrónico.

A su vez, LegalXML, produce estándares para la presentación de citas, transcripciones. Dentro de sus miembros y contribuidores, se encuentran entre otros: Africa i-Parliaments Action Plan (UN/DESA), CATCert Agencia Catalana de Certificación, la Biblioteca del Congreso Nacional de Chile, Directorate of Labour and Welfare Norway (Dirección General de Trabajo y Bienestar Social de Noruega), el Parlamento Europeo, Netherlands Tax and Customs Administration (Administración Aduanera y Tributaria) (Belastingdienst), University of Bologna (Universidad de Bolonia) CIRSIFID, el Parlamento de Uruguay y el National Center for State Courts (Centro Nacional para Tribunales Estatales).

De igual forma, como ya se mencionó, LegalXML, se encuentra vinculada con OASIS, consorcio sin ánimo de lucro, cuyos fines son el impulso del desarrollo, la convergencia y adopción de estándares abiertos para la sociedad de la información mundial. Los estándares creados por OASIS, son abiertos, estimulan la innovación, buscan la disminución de costos y protegen el derecho de la libre elección de la tecnología.

Ahora bien, a partir del año 2001, una entidad que inicia su trabajo con XML, es la House of Representatives (Cámara de Representantes) de Estados Unidos, que pretende proporcionar información acerca de los procesos en curso dentro de la misma entidad con la utilización del lenguaje XML. Es de anotar, que si bien el proyecto ya se encuentra en curso, no ha culminado su etapa de desarrollo por lo cual incluso dentro de su página web²⁷⁸, señala que este se puede extender en el tiempo. Sin embargo, es clara cuál es su tecnología base: “Las estructuras de datos para XML se materializan en documentos DTD (Data Type Definition) o a través de esquemas XML. En el ámbito legal deben utilizarse estas estructuras de datos para documentos tales como sentencias judiciales, normas jurídicas y contratos, describiéndose bien por conducto de DTD o Esquemas XML”²⁷⁹.

Al mismo tiempo, desde febrero de 1998, surge LEXML²⁸⁰, como la red europea que al igual que LegisXML, pretende alcanzar el intercambio automático de información jurídica y su estandarización. La descripción de los datos, que para los documentos jurídicos resulta indispensable, se implementa por la descripción a través de DTD o Esquemas XML. Es de anotar que LEXML, se encuentra en Alemania, Suecia, Holanda, Italia y España. En igual sentido, LeXML, se encuentra encargada del desarrollo de estándares XML aplicados al ámbito jurídico para Europa.

“La aplicación de las tecnologías XML a los documentos jurídicos en los distintos países europeos ha originado iniciativas y proyectos de forma

²⁷⁸ Xml.house.gov/ (n.d.). Consultada en mayo 11, 2012. Xml.house: <http://xml.house.gov>

²⁷⁹ Lexml.es. (n.d.). Consultada en mayo 11, 2012. Lexml.es: <http://www.uv.es/lexml/>

²⁸⁰ Ibídem.

continuada. La Unión Europea ha encabezado la apuesta por un marco de interoperabilidad para atender los servicios derivados del e-gobierno desarrollando numerosos programas, como el reciente Semantic Interoperability Strategy / XML Clearinghouse.

Queremos subrayar en este punto una serie de proyectos enmarcados dentro de las políticas generales de información que diversos países europeos han puesto en marcha en los últimos años. En el Reino Unido se lanza en 2000 un plan de modernización de las administraciones públicas que integra sus sistemas de información con el programa e-GIF (e-Government Interoperability Framework), tomando como base tecnológica XML. En la actualidad, el portal UK GovTalk (<http://www.govtalk.gov.uk>) pretende desarrollar políticas y estándares para la e-administración y ponerlas a disposición de los sectores público y privado”²⁸¹.

2.15. Bancos de datos jurídicos en Internet

En lo relacionado con la informática jurídica documental, dentro del periodo que comprende el surgimiento de Internet a la actualidad, como examinaremos a continuación, se puede dilucidar un crecimiento exponencial de la organización y sistematización de los contenidos jurídicos distribuidos en nuestro país y en el mundo.

²⁸¹ ALVITE DÍEZ, M. L. (2009). op.cit. p. 5.

No obstante, los bancos de datos que se encuentran en Internet, tienen diferentes matices y variables en atención al Estado del cual se haga referencia. En el caso particular de Colombia, como examinaremos en líneas posteriores, podremos determinar clara y categóricamente que los bancos de datos jurídicos que se ofrecen en Internet, provienen en mayor cantidad de casos del sector privado que no del sector público.

Contrario a lo anterior, en países de tradición igualmente continental como Francia o Bélgica; por mencionar algunos, el gobierno concentra esfuerzos para transmitir la información jurídica de forma libre, sin restricción alguna de manera que cualquier habitante del planeta pueda acceder y conocer de manera organizada en el inmenso mundo del ordenamiento jurídico, aquello que se encuentra vigente y en especial las nuevas actualidades del derecho, hecho que permite que el ciudadano conozca las leyes y el operador jurídico pueda realizar su trabajo con las herramientas adecuadas.

Ahora bien, al igual que sucede en Colombia, en España la informática jurídica documental es proporcionada en mayor medida por los empresarios privados, en este sentido se ha mencionado lo siguiente:

“La distribución de contenidos jurídicos en nuestro país cuenta con una participación mayoritaria de la iniciativa privada, aspecto este inusual con respecto a la industria de las bases de datos en España. Esta peculiaridad guarda una vinculación directa con la importancia histórica de las empresas privadas en la publicación de repertorios jurídicos, pero ahora han de competir, manteniendo sus objetivos comerciales, con fuentes de información

de nuevo cuño, que ofrecen en la Red, en muchos casos, información gratuita. Un nuevo contexto en el que se precisa demostrar un alto valor añadido que singularice contenidos de pago, en apariencia semejantes a los disponibles sin coste alguno”²⁸².

Para concluir este segundo capítulo, en aras de examinar el panorama mundial de los bancos de datos jurídicos que se encuentran en Internet, analizaremos a continuación en primer lugar los bancos de datos jurídicos de la informática jurídica documental del pasado, que se encuentran vigentes en la actualidad y de forma paulatina, cuales han surgido en esta nueva era de la sociedad de la información.

2.15.1. Bélgica

En este país, continúa vigente el CREDOC (Centre de Documentation Automatique du Droit). Que surgió en el año de 1967, con su presentación en idioma francés y neerlandés. En la actualidad, cuenta con la página web www.credoc.be. De igual manera, CREDOC continúa propendiendo por el mejoramiento de los servicios notariales y busca corregir aspectos como la inseguridad jurídica.

En igual sentido, CREDOC cuenta con conocimientos acerca de la automatización del derecho y centra su atención en el desarrollo de las bases de datos contentivas de las disposiciones notariales y en la automatización de la labor notarial. Por tal motivo, son abanderados en la informática jurídica del notariado belga.

²⁸² Ibidem. p. 2.

Al mismo tiempo, CREDOC ofrece servicios de redes, seguridad informática, aplicaciones informáticas complejas para los notarios, traducciones de documentos jurídicos, legalización de los mismos y mantenimiento de las bases de datos. Por lo cual, como se puede observar, ofrece la posibilidad al operador jurídico de analizar y acceder a la información de forma organizada, ágil y efectiva. Además otorga las herramientas necesarias para facilitar y brindar mejor funcionamiento en el ámbito del derecho notarial.

También localizamos en Bélgica, el servicio público federal de justicia, que suministra una página desde la cual se accede alfabéticamente a los distintos temas de interés público, relacionados con el tema de justicia²⁸³.

Asimismo, encontramos el sistema de información JURIDAT, corresponde al portal de la judicatura de Bélgica. En dicho país es posible acceder a la legislación, jurisprudencia y boletín oficial Moniteur belge, mediante dicho portal. El acceso al banco de datos es gratuito²⁸⁴.

La consulta y recuperación de la jurisprudencia, se logra mediante la consulta de los tesauros, que se encuentran organizados en forma arborescente, por cada una de las ramas del derecho. Adicionalmente es posible conocer la jurisdicción, la localización y el idioma, según el tribunal y los tres idiomas oficiales. Igualmente,

²⁸³ SERVICE PUBLIC FÉDÉRAL JUSTICE. (2014). Consultada en julio 30, 2014. Service Public Fédéral Justice: <http://justice.belgium.be/fr/>

²⁸⁴ THE EUROPEAN E-JUSTICE PORTAL. (2014). Consultada en julio 30, 2014. The European e-Justice Portal: <https://e-justice.europa.eu/home.do?plang=en&action=home>

la búsqueda se puede hacer por fecha y se puede hacer uso de los conectores. Como ejemplo podemos mencionar el Portal del Poder Judicial de Bélgica. Juridat²⁸⁵.

La consulta de la legislación se efectúa por la naturaleza del documento, si es un código, constitución, reglamentos, tratados. Igualmente, se puede realizar por número, fecha de promulgación o publicación, palabras con el conector (Y), y búsqueda sobre texto o título²⁸⁶.

2.15.2. Francia

La evolución de la informática jurídica en Francia es avanzada, existen bases de datos jurídicas de origen gubernamental y otras de origen privado. Igualmente, las bases de datos jurídicas cuyo nacimiento es público cuentan con un acceso ilimitado para toda la sociedad de la información en internet.

En Francia encontramos los bancos de datos jurídicos de acceso gratuito, siendo la más utilizada y emblemática la base de datos jurídica legifrance²⁸⁷, ella contiene los códigos en vigor, las leyes y reglamentos, así como la jurisprudencia constitucional, administrativa y judicial y las convenciones colectivas. La búsqueda puede ser simple o experta.

De igual forma, la página ofrece acceso al diario oficial, al cual es posible suscribirse de forma gratuita, obtenerlo diariamente en el correo electrónico del usuario. Legifrance, a su vez ofrece la actualidad jurídica sobre los proyectos legislativos, la

²⁸⁵ LE PORTAIL DU POUVOIR JUDICIAIRE DE BELGIQUE. (n.d.). Consultada en marzo 13, 2015. Juridat: <http://www.juridat.be>

²⁸⁶ Ibídem.

²⁸⁷ LEGIFRANCE. (n.d.). Consultada en marzo 13, 2015. Legifrance: <http://www.legifrance.gouv.fr/>

actualidad de la Unión Europea y la actualidad internacional. La página se encuentra en francés, no obstante, ofrece traducciones en árabe, alemán, inglés, español, italiano, entre otros.

En cuanto a la técnica de organización de la información, son los tesauros formados por sus respectivos descriptores los que forman la base para la creación de las bases de datos y su posterior búsqueda. Es posible realizar búsquedas sobre un tema y encontrar resultados de diferentes áreas del derecho en el inicio de la página. Es una página web muy completa que facilita el conocimiento del derecho a los internautas al ofrecer su servicio sin restricción alguna.

Dentro de los sitios que ofrecen bases de datos de acceso gratuito, también se encuentran los sitios de internet del parlamento francés, a saber: la Asamblea Nacional²⁸⁸ y el Senado²⁸⁹. Sobre estos sitios se realiza la actualización del conjunto de trabajos y documentos del parlamento, así como las noticias biográficas de los diputados y senadores.

Es importante mencionar que la rama judicial realiza las publicaciones de sus decisiones en los sitios de Internet de la Corte de Casación²⁹⁰, el Consejo Constitucional²⁹¹, del Consejo de Estado y de la Jurisdicción Administrativa²⁹² y del Tribunal de Conflicto²⁹³, entre otros.

²⁸⁸ ASSEMBLÉE NATIONALE FRANCE. (n.d.). Consultada en marzo 13, 2015. ASSEMBLÉE: <http://www.assemblee-nationale.fr/index.asp>

²⁸⁹ SENAT. (n.d.). Consultada en marzo 13, 2015. Senat: <http://www.senat.fr/>

²⁹⁰ COUR DE CASSATION. (n.d.). Consultada en marzo 13, 2015. Cour de cassation. Jurisprudence: https://www.courdecassation.fr/jurisprudence_2/

²⁹¹ CONSEIL CONSTITUTIONNEL. (n.d.). Consultada en marzo 13, 2015. Conseil constitutionnel. Les décisions: <http://www.conseil-constitutionnel.fr/conseil-constitutionnel/francais/les-decisions/les-decisions.95486.html>

También, se ubican Bases de datos jurídicas con acceso remunerado, dentro de las cuales está la base de datos jurídica de Dalloz²⁹⁴ que contiene la jurisprudencia de las altas cortes y de los tribunales de instancia, así como los códigos y textos no codificados. Dentro de las bases de datos jurídicas de Dalloz se encuentran igualmente las obras jurídicas de la doctrina francesa y en especial revistas dedicadas al análisis de la jurisprudencia y a los temas que suscitan controversias jurídicas en torno a las diferentes ramas del derecho como: el derecho de familia, penal, civil, comercial, de comercio electrónico, derecho privado internacional, derechos de autor, entre muchos otros. Dalloz cuenta con la participación de académicos destacados y las opciones de búsqueda son variadas, es posible introducir un descriptor y obtener un resultado de búsqueda sobre diferentes áreas del derecho o sobre alguna en particular. Estas bases de datos jurídicas son ampliamente reconocidas en el ámbito jurídico francés.

Dentro de las bases de datos jurídicas que exigen inscripción, se halla igualmente lexisnexis jurisclasseur²⁹⁵, ese sitio ofrece búsquedas con la utilización de operadores booleanos. El sitio ofrece sus revistas jurídicas, los códigos y leyes, enciclopedias,

²⁹² « Cette base de jurisprudence donne accès à plus de 230 000 documents: décisions et avis contentieux du Conseil d'État et des cours administratives d'appel, analyses de ces décisions et avis ayant été retenus pour leur apport à la jurisprudence, ainsi qu'une sélection de conclusions de rapporteurs publics ».

Traducción propia: Esta base de jurisprudencia da acceso a mas de 230.000 documentos : decisiones y conceptos contenciosos del Consejo de Estado y de las cortes administrativas de apelación, análisis de estas decisiones y conceptos que han sido retenidos por su aporte con la jurisprudencia, así como una selección de conclusiones de los relatores públicos. LE CONSEIL D'ÉTAT ET LA JURIDICTION ADMINISTRATIVE. (n.d.). Consultada en marzo 13, 2015. Le Conseil d'État et la Juridiction Administrative: <http://www.conseil-etat.fr/Decisions-Avis-Publications/Decisions/ArianeWeb>

²⁹³ TRIBUNAL DES CONFLICTS. (n.d.). Consultada en marzo 13, 2015. Decisions: http://www.tribunal-conflits.fr/decisions_2014.html

²⁹⁴ DALLOZ.FR. (2015). Consultada en marzo 13, 2015. Dalloz.fr: <http://www.dalloz.fr/>

²⁹⁵ LEXISNEXIS JURISCLASSEUR. (2015). Consultada en marzo 13, 2015. LexisNexis JurisClasseur: <http://www.lexisnexis.com/fr/droit/>

actualidades jurídicas, toda la jurisprudencia, el diario oficial, la reglamentación europea y las convenciones colectivas.

2.15.3. Reino Unido

En el Reino Unido, los primeros experimentos conducidos por Colin Tapper se remontan al año 1963. Entre 1968 y 1969 se diseñó el SRI legal STATUS en el U.K. Atomic Energy Authority en Harwell. Sin embargo, la puesta en marcha de servicios comerciales no llegaría hasta el año 1978, momento en el que la editorial Butterworths anunció su acuerdo con Mead Data Central para ofrecer el sistema LEXIS²⁹⁶.

Los bancos de datos jurídicos existentes en el Reino Unido tienen un valor primordial en la práctica de los abogados solicitors²⁹⁷ y barristers²⁹⁸ en el caso de Inglaterra, Gales e Irlanda del Norte, y de Lawyers y advocates, en el caso de Escocia²⁹⁹.

En el Reino Unido existen varios sitios de información jurídica, entre los que encontramos el sitio web de la legislación³⁰⁰, que entrega el texto completo de la

²⁹⁶ ALVITE DÍEZ, M. L. (2003). op.cit. p. 192.

²⁹⁷ Los abogados solicitors son los que se encargan de la preparación legal del caso, que implica el análisis de los hechos y la identificación de los documentos relevantes. Asesoran a sus clientes sobre la normativa aplicable y sobre modos alternativos de resolución de conflictos, preparan conclusiones y actúan como enlace con el abogado de la contraparte, son el principal contacto del cliente.

²⁹⁸ Los barristers son una de las dos categorías de abogados que existen en Inglaterra, Escocia (en donde se utiliza el término "*advocate*") y su principal función es representar como mandatarios a los litigantes ante los tribunales.

²⁹⁹ *The Bodleian Law Library*. (2014). Consultada en agosto 8, 2014. Bases: <http://www.bodleian.ox.ac.uk/law>

³⁰⁰ *UK. Legislation*. (n.d.). Consultada en agosto 5, 2014. UK. Legislation: <http://www.legislation.gov.uk/>

legislación primaria, promulgada por el Parlamento del Reino Unido, el Parlamento de Escocia, la Asamblea de Irlanda del Norte y la Asamblea Nacional de Gales, junto con toda la legislación secundaria aplicable al Reino Unido. El acceso a la información mencionada, es gratuito. En la base de datos legislativa Statute Law Database, se encuentra la legislación primaria revisada de todos los lugares del Reino Unido, desde 1235 hasta la actualidad y su acceso es gratuito³⁰¹.

También encontramos la base de datos del británico e irlandés Instituto de Información Legal bailii, que suministra los documentos legales primarios británicos e irlandeses, la jurisprudencia de la Unión Europea, Comisión de Derecho informa, y otro material británica e irlandesa relacionada con la ley, de utilización gratuita y en un solo lugar en Internet. En agosto de 2012, bailii incluyó 90 bases de datos que cubren 7 jurisdicciones. El sistema contiene alrededor de 36 gigabytes de materiales legales y alrededor de 297.513 documentos posibles de búsqueda³⁰².

Por la naturaleza del sistema jurídico anglosajón en el Reino Unido la indexación de la jurisprudencia ha tenido un gran desarrollo, ya que dentro de la competencia profesional, es muy importante la capacidad del jurista para conseguir precedentes de alta jerarquía judicial. Bailii contiene una recopilación de las sentencias del Tribunal Superior de Justicia desde 1879, del Tribunal Superior de lo Penal desde 1914, de los juzgados de primera instancia desde 1998 y de la Cámara de los Lores desde 1838. Gran parte de la jurisprudencia de Escocia se encuentra a disposición del público. En esta página encontramos sentencias de diversos Tribunales del Reino Unido como,

³⁰¹ PORTAL EUROPEO DE E-JUSTICIA. (n.d.). Consultada en agosto 5, 2014. Portal Europeo de e-Justicia: <https://e-justice.europa.eu/home.do>

³⁰² BRITISH AND IRISH LEGAL INFORMATION INSTITUTE. (2010, Noviembre 25). Consultada en agosto 20, 2014. Bases de Datos: <http://www.bailii.org/databases.html>

Tribunales Especiales de Nivel Superior (Apelación Administrativa, Hacienda e Impuestos y Propiedad Inmobiliaria), el Tribunal Especial de Asilo e Inmigración del Reino Unido y Tribunal Especial de Información del Reino Unido, incluida la Comisión de Apelación en materia de Seguridad Nacional, entre otros³⁰³.

Bailii se encuentra legalmente constituida en el Reino Unido como una compañía limitada por garantía y como un fondo de caridad. Se encuentra hospedado en el Reino Unido e Irlanda, por el Instituto de Estudios Jurídicos Avanzados, Londres y la Facultad de Derecho, Universidad de Cork³⁰⁴.

Las bases de datos se pueden consultar por palabra clave o navegar en ellas por fecha o lista de títulos, presentados en orden alfabético, contienen material reciente, pero la cobertura varía y tanto los antiguos como los nuevos contenidos se están agregando en forma permanente³⁰⁵.

En la página del Parlamento del Reino Unido³⁰⁶, se encuentra el archivo de las sentencias promulgadas por la Cámara de los Lores, desde el 14 de noviembre de 1996 hasta 30 de julio de 2009, ya que posterior a esta última fecha, la Supreme Court of the United Kingdom, asumió la jurisdicción sobre cuestiones de derecho para todos los casos de derecho civil en el Reino Unido y todos los casos penales en

³⁰³ *Ibidem*.

³⁰⁴ *Ibidem*.

³⁰⁵ *Ibidem*.

³⁰⁶ THE PARLIAMENT OF THE UNITED KINGDOM OF GREAT BRITAIN AND NORTHERN IRELAND. (n.d.). Consultada en marzo 14, 2015. The Parliament of the United Kingdom of Great Britain and Northern Ireland: <http://www.parliament.uk/>

Inglaterra y Gales e Irlanda del Norte. En su página web³⁰⁷, relaciona las providencias proferidas por este órgano, en formato PDF y con acceso gratuito.

La base de datos Westlaw Reino Unido, diseñada para el entorno digital, permite el acceso a la jurisprudencia de algunos de los más renombrados serie de informes del Reino Unido, con más de cuatrocientos mil informes de casos y las transcripciones; legislación consolidada desde 1267, que se actualiza diariamente; noticias; revistas jurídicas, más de ciento diez revistas, con acceso a los resúmenes de cada artículo en cada revista jurídica; comentarios; alertas actuales de sensibilización y materiales jurídicos de la UE³⁰⁸.

Westlaw Reino Unido, es una base de datos de acceso restringido a suscritos, producida por Thomson Reuters. La búsqueda de la información se hace mediante dos formas básicas: Insight, que cuenta con un índice navegable de temas jurídicos, y por índice de temas, que conecta con definiciones que posteriormente enlaza con las autoridades relevantes de dicho tema. Adicionalmente, Westlaw UK, tiene un buscador común, en el que se tiene acceso por término legal o por conducto de un sistema de reconocimiento de lenguaje natural, creando filtros por tipo de material (legislación, jurisprudencia, etc.), fecha, conectores y de búsqueda sobre búsqueda o búsqueda dentro de los resultados³⁰⁹.

Del mismo modo ubicamos en el Reino Unido a Lexis Library, anteriormente denominada Butterworths. Lexis Library es una base de datos jurídica que

³⁰⁷ THE SUPREME COURT OF UNITED KINGDOM. (2015). Consultada en marzo 14, 2015. Decided cases: <https://www.supremecourt.uk/index.html>

³⁰⁸ WESTLAW UK. (2014). Consultada en agosto 21, 2014. Westlaw UK: <http://legalresearch.westlaw.co.uk/>

³⁰⁹ *Ibíd.*

proporciona acceso a la jurisprudencia en texto completo, legislación a partir de 1267 en adelante, incluye el texto completo, en su versión modificada de todos los actos que están actualmente en vigor y la legislación que se ha promulgado recientemente, pero aún no está en vigor y los artículos de revistas seleccionadas desde el Reino Unido, la Unión Europea, Estados Unidos y otras jurisdicciones seleccionadas en todo el mundo³¹⁰.

2.15.4. Canadá

En este país, nos encontramos con varios centros dedicados a la informática jurídica documental, dentro de ellos, un banco que se encontraba desde el pasado de la informática jurídica, este es el Centro de Documentación de la Universidad de Laval en Montreal.

Dentro de los contenidos del Centro de Documentación, reposan entre otros, los proyectos de ley para ser sometidos a consideración del gobierno federal y del gobierno de Quebec. Así mismo, se incluyen dentro de sus bases de datos los boletines, convenciones colectivas, el reporte del congreso y las tesis doctorales de la Universidad de Laval.

El Centro de Documentación de la Universidad de Laval³¹¹, presenta a disposición del usuario una colección rica e inédita sobre el derecho laboral. Además existe un

³¹⁰ UNIVERSITY OF LONDON. SCHOOL OF ORIENTAL AND AFRICAN STUDIES. (2007). Consultada en agosto 25, 2014. Bases de datos: <http://www.soas.ac.uk/library/subjects/law/databases/lexisnexis-butterworths/>

³¹¹ UNIVERSITÉ LAVAL. FACULTE DES SCIENCES SOCIALES. Département des Relations Industrielles. (2015). Consultada en marzo 14, 2015. Documentos: <http://www.rlt.ulaval.ca/documents/accueil.asp>

compendio extenso de periódicos, documentos sindicales, gubernamentales como ya se mencionó y las ediciones de la revista de Relaciones Laborales de la misma Universidad. Además de tesis, memorias, ensayos en relaciones industriales y un sin número de publicaciones de docentes a las cuales se cuenta con la posibilidad de acceder desde cualquier parte del mundo.

Finalmente, es importante mencionar que dentro de este centro de documentación, como bien se mencionó en el pasado de la informática jurídica, se incorporan herramientas para evitar la repetición de trabajos y realizar intercambios de microfilms indexados con otras universidades, de manera que además de incluir bases de datos realiza una labor de cooperación internacional con el suministro de esta información.

Por otra parte, en la actualidad encontramos que en Canadá, uno de los más grandes difusores de la informática jurídica documental en Internet es LEXUM³¹², empresa interdisciplinar dedicada a estos menesteres desde el año 2009, fecha de su fundación en Quebec.

Esta empresa prestadora de servicios en informática jurídica, ofrece asistencia tanto a los productores como a los usuarios de la información del ordenamiento jurídico. Además de ofrecer sus productos, Lexum ofrece soluciones personalizadas que atienden a las necesidades del usuario de la información jurídica.

³¹² LEXUM. (2015). Consultada en marzo, 2015. LexUM: <https://lexum.com/fr/>

De igual forma, Lexum se encuentra compuesto por un equipo multidisciplinario de más de 40 expertos en informática, derecho y en general de todas las ciencias de la información. Hoy en día, es considerada la más importante empresa canadiense de información jurídica publicada en Internet.

Desde sus inicios en el año de 1993, dentro del marco de colaboración con la Corte Suprema de Canadá, cuyo objeto consistía en la publicación de la totalidad de sus publicaciones en Internet, Lexum ha incluido gran variedad de servicios, entre los cuales podemos mencionar: EL servicio de consultas, el diseño, desarrollo, alojamiento y soporte a largo plazo a instituciones jurídicas canadienses, organizaciones internacionales como el alto comisionado para los derechos del hombre de Naciones Unidas, la Unión Internacional de las telecomunicaciones y el Centro de Comercio Internacional.

Es importante reconocer que Lexum es un miembro fundador del movimiento por el acceso libre al derecho, movimiento que reagrupa más de treinta organizaciones a escala mundial y que propende por lograr optimizar y garantizar el derecho a libre acceso de la información jurídica para todos los ciudadanos en cada país miembro. Valga mencionar que este movimiento es de vital importancia y se encuentra a la vanguardia de lo que hoy en día reposa en Internet en lo concerniente a la informática jurídica. Por tal motivo al finalizar esta segunda parte terminaremos con una reflexión a propósito de este importante tema.

En consecuencia, Lexum es reconocido en Canadá como el diseñador y operador del sitio web CanLII³¹³, que se encuentra posicionado como el principal recurso de informática jurídica en Canadá, este sitio ofrece la posibilidad de acceder de forma gratuita a todas las leyes y ordenamiento jurídico en general de todas las jurisdicciones canadienses, al igual que a las decisiones judiciales provenientes de las cortes y tribunales canadienses.

2.15.5. España

En España, dentro de la informática jurídica del pasado, aludimos a varios sistemas de tratamiento y recuperación de la información, dentro de ellos Andrid del Ministerio de Justicia, Ibertrad que dependía del Ministerio Español de Asuntos Exteriores y Sinade que pertenecía al Colegio de Abogados de Barcelona.

Además de los bancos de datos jurídicos citados en el pasado, en España la informática jurídica se ha caracterizado por una amplia expansión. Podemos encontrar en primera instancia bases de datos jurídicas comerciales, en segundo lugar y no menos importante, existen diferentes portales institucionales, dentro de los cuales se encuentran los sistemas parlamentarios, los diarios oficiales y el ámbito judicial. Finalmente dentro de la informática jurídica actual, podemos constatar la materialización de proyectos de investigación, tales como XML legislativo.

Corresponde examinar en lo sucesivo cada uno de los componentes de la informática jurídica española ya anteriormente enumerados. En primer lugar se encuentran las

³¹³ INSTITUT CANADIEN D'INFORMATION JURIDIQUE. CANLII. (n.d.). Consultada en marzo 14, 2015. <http://www.canlii.org/>

bases de datos jurídicas comerciales, no obstante, antes de entrar a explicarlas, es menester mencionar que al igual que Canadá, la tendencia en España en lo relacionado con la consulta de sus contenidos es hacia la gratuidad de las mismas.

Una vez mencionado lo anterior, podremos citar en primer lugar el portal jurídico Derecho.com³¹⁴, cuya fecha de creación data de 1994, es dirigido por jurisperitos, desde el año 2002 ha comenzado a trabajar con el formato XML (antes ya explicado en anteriores acápites) y su sistema funciona en Linux. En su página de inicio ofrece la posibilidad al usuario de informarse sobre las noticias jurídicas más relevantes del país.

De igual forma para hacer uso de los bancos de datos es necesario realizar una suscripción de manera que su acceso es de contenido limitado, dentro de sus obras se encuentra una de contenido general que cuenta con jurisprudencia, legislación, doctrina actualizada y ofrece acceso ilimitado en tiempo y espacio.

En cuanto a las obras especializadas, que son más profundas y rigurosas en cuanto a sus contenidos, el derecho.com ofrece a sus usuarios bancos de datos sobre la contratación inmobiliaria, la propiedad horizontal y los derecho reales, derecho mercantil, derecho fiscal, social, tráfico; es decir derecho de circulación y seguro de vehículos, el derecho de urbanismo, la gestión administrativa, el banco de datos contable, banco de derecho de familia.

Igualmente, a los usuarios que requieren información aún más especializada, esta plataforma ofrece información de la administración local, para efectos de permitir el

³¹⁴ PORTAL JURÍDICO EL DERECHO. (n.d.). Consultada en marzo 14, 2015. Portal Jurídico El Derecho: <http://www.elderecho.com/>

desarrollo de una adecuada actividad en una entidad local o un territorio concreto, este banco de datos es denominado el derecho local.

Otra particularidad de estas bases de datos es la posibilidad de consultar los tesauros dentro de las áreas temáticas de derecho administrativo, civil, comunitario, constitucional, internacional, fiscal, mercantil, penal, social-laboral, tecnología y contable. Además ofrece la información en consideración a la profesión del operador jurídico y del área geográfica de España, que desea consultar.

De otra parte encontramos la editorial Lex Nova³¹⁵, que ha sido desarrollada con el grupo de investigación de Recuperación de Información y Bibliotecas Digitales de la Facultad de Informática de la Universidad de Valladolid, surge en 1953 como editorial especializada en el derecho laboral, es de resaltar que esta editorial hace uso de los metadatos y trabaja en XML.

Lex Nova, a su vez cuenta con bancos de datos en las áreas del derecho administrativo, civil, migratorios, familia, fiscal, laboral, mercantil, penal, prevención, sanitario, tráfico entre otras áreas como contabilidad y financiero. En materia de documentos jurídicos este portal ofrece a sus usuarios jurisprudencia, legislación, convenios, minutas y casos jurídicos prácticos, doctrina, gacetas informativas.

³¹⁵ LEX NOVA. (n.d.). Consultada en marzo 14, 2015. Lex Nova: <http://www.lexnova.es/>

En lo relacionado con el acceso a la información Lex Nova ofrece acceso libre a la legislación, la jurisprudencia y la práctica, sin embargo para realizar consultas especializadas dentro de sus diferentes productos se requiere una suscripción, este portal jurídico ofrece diferentes formatos a sus suscriptores, existe la posibilidad de hacer uso del servicio de almacenamiento, manipulación y envío de bibliografía para el operador jurídico.

Reposa igualmente en Internet, la base de datos jurídica denominada Tirant³¹⁶, dirigido por la empresa ISOCO, esta base de datos surge en el año 2002 y tienen como particularidad el uso de “técnicas de inteligencia artificial para tratamiento de textos, técnicas avanzadas de búsqueda, seguridad en las conexiones, autenticación de usuarios y XML. El formato XML se usa para introducir y clasificar de forma automática los documentos en el sistema, permitiendo una conveniente separación entre la capa de datos, la lógica de la aplicación y el sistema de visualización. Además, emplean vocabularios controlados: listas de palabras clave, tesauros, metadatos y ontologías”³¹⁷.

El acceso a sus bases se realiza únicamente vía Internet, existe un acceso libre restringido y sus bancos de datos contienen información sobre derecho penal, público, financiero, tributario, privado, procesal civil, laboral, jurisprudencia de Aragón, de la Comunidad Valenciana, Ourense, Cataluña, Guadalajara, revistas jurídico y ordenanzas locales. De igual forma dentro de la documentación jurídica

³¹⁶ TIRANT LO BLANCH. (n.d.). Consultada en marzo 14, 2015. Tirant Lo Blanch: <http://www.tirantonline.com/tol/>

³¹⁷ ALVITE DÍEZ, M. L. (2009). op.cit. p. 10.

ofrece legislación, jurisprudencia, doctrina, bibliografía de todas las editoriales, formularios, modelos, esquemas de derecho procesal, sustantivo y consultas.

Sus particularidades son varias, dentro de ellas es posible mencionar que además de organizar la información esta plataforma web la integra, asunto que no es frecuentemente realizado en nivel avanzado por las otras bases de datos jurídicas que reposan en Internet, la integración de la información ofrece la posibilidad al usuario de observar desde diferentes perspectivas su búsqueda y obtener un conocimiento íntegro de la información que se encuentra organizada por sistemas de navegación.

Los sistemas de búsqueda de esta plataforma son dos a saber: los primeros, son los buscadores determinados para cada clase de documento jurídico y el segundo de ellos es el buscador general que actúa sobre toda la base de datos, ellos en su conjunto otorgan una navegación y búsqueda intuitiva. Sus actualizaciones se realizan diariamente y las actualizaciones son enviadas a cada usuario, no obstante pueden ser igualmente consultadas en la página web.

Una vez explicadas las bases de datos comerciales, hablaremos ahora de los portales institucionales, para esto es importante mencionar que dentro de ellos reposan a saber: Los sistemas parlamentarios, los diarios oficiales, las plataformas de e-gobierno y el ámbito judicial. En este orden de ideas explicaremos en primera instancia los sistemas parlamentarios.

Dentro de los sistemas parlamentarios, el senado ha implementado el tesoro denominado Eurovoc, con el fin de realizar búsquedas semánticas en cada una de las

ocho bases de datos jurídicas ofrecidas al público. Es de señalar, que dentro de sus versiones se “ha empleado la edición 4 del tesoro Eurovoc, un desarrollo propio del Senado y la Relación de municipios y códigos por provincias publicada por el Instituto Nacional de Estadística”³¹⁸. De esta manera, los resultados de la búsqueda se obtendrán desde las diferentes perspectivas en atención a su semántica, sinonimia, conexidad, relaciones asociativas y de jerarquía. Es de anotar que el desarrollo de Eurovoc por el senado español, forma parte de una plataforma de informática jurídica de la Unión Europea, de forma que la explicación de su funcionamiento y alcance tendrá lugar dentro del acápite de los bancos de datos jurídicos de la Unión Europea.

Para lo que nos concierne en este punto, es relevante mencionar que Eurovoc es un tesoro multilingüe que contiene información documental de las instituciones europeas en los ámbitos comunitarios y de los aspectos internos de cada país miembro, en especial dentro de lo relacionado con la legislación, de ahí que el senado español participe en su formación.

El fin principal de la realización del tesoro consiste en aportar la información a los servicios responsables de la gestión y difusión, de manera que el tesoro se convierte en el instrumento por excelencia para la indexación coherente para proporcionar búsquedas satisfactorias a los usuarios.

En segundo lugar, dentro de los sistemas parlamentarios, encontramos que las Cortes de Aragón utilizan la aplicación Lotus notes, este portal tiene origen en el año de

³¹⁸ SENADO ESPAÑOL. (n.d.). Consultada en marzo 14, 2015. Senado Español: <http://www.senado.es/brsweb/notastes.html/>

2006 y trabaja actualmente en un nuevo desarrollo, para emprender búsquedas ofrece ficheros en formato XML.

Igualmente, la Junta General del Principado de Asturias³¹⁹, en el año 2006, implementó el formato XML en su portal, este hace uso de los metadatos, los tesauros y los descriptores. De igual forma, el Parlamento de Baleares³²⁰, desde el año 2001 utiliza el formato XML con la edición tercera de Eurovoc que se encuentra aplicada por el senado español y traducido al idioma catalán.

Como las anteriores plataformas, los portales del parlamento de Cantabria³²¹, las Cortes de Castilla y León³²² y el Parlamento del País Vasco³²³, hacen uso del formato XML, el primero de ellos utiliza metadatos y plataformas de palabras clave, además es desarrollado por EUREKA mediante el software denominado Nhibernate, en segundo lugar las Cortes de Castilla y León incluyen dentro de sus bases de datos fragmentos de video de las sesiones parlamentarias en formato XML y para finalizar el País Vasco utiliza el software a medida y Eurovoc para realizar el control del vocabulario.

Ahora bien, en lo concerniente a los diarios oficiales, existen dos sistemas conocidos por sus avances tecnológicos, estos son el del Principado de Asturias y el del Boletín

³¹⁹ JUNTA GENERAL DEL PRINCIPADO DE ASTURIAS. (n.d.). Consultada en marzo 14, 2015. Junta General del Principado de Asturias: <http://www.jgpa.es/>

³²⁰ PARLAMENT DE LES ILLES BALEARS. (n.d.). Consultada en marzo 14, 2015. Parlament de les Illes Balears: <http://www.parlamentib.es/Home/Default.aspx>

³²¹ PARLAMENTO DE CANTABRIA. (n.d.). Consultada en marzo 14, 2015. Parlamento de Cantabria: <http://www.parlamento-cantabria.es/>

³²² CORTES DE CASTILLA Y LEÓN. (n.d.). Consultada en marzo 14, 2015. <http://www.ccyl.es/export/sites/ccyl/index.html>

³²³ PARLAMENTO VASCO. (n.d.). Consultada en marzo 14, 2015. <http://www.parlamentovasco.euskolegebiltzarra.org/es>

Oficial de la Región de Murcia. En cuanto al primero, el sistema es denominado en sus siglas (SIDRA) Sistema de Información Documental en Red de Asturias³²⁴, este sistema se encuentra operando en formato XML, además de presentar su plataforma en Internet, este diario oficial publica de forma semestral sus índices en DVD y utiliza metadatos, tesauros y descriptores con el fin de organizar las búsquedas.

En lo relacionado con el Boletín Oficial de la Región de Murcia, el formato web se encuentra disponible desde 1996, en la actualidad usa el formato XML utilizado en la web 2.0, al igual que el boletín oficial del Principado de Asturias presenta una edición en CD-ROM anualmente y utiliza descriptores, metadatos y tesauros para la organización y facilitación de la búsqueda de la información.

Ahora bien, en cuanto al Boletín Oficial de España³²⁵, este inicia en el año 2002 su desarrollo, utiliza de igual forma los componentes fundamentales para un buen avance de la informática jurídica documental, es decir los tesauros y descriptores. Su acceso es gratuito para los usuarios en la web y en la actualidad no presenta algún otro formato, por lo cual esta se constituye como la única forma de consulta.

Para finalizar, se presenta la situación del ámbito judicial de las bases de datos en España, que gracias a la Recomendación 2001/3 del Comité de Ministros a los Estados miembros de la Unión Europea que considera como responsabilidad pública la distribución de información jurídica, España, al igual que los demás países

³²⁴ GOBIERNO DEL PRINCIPADO DE ASTURIAS. (n.d.). Consultada en diciembre 28, 2012. <https://sede.asturias.es/portal/site/Asturias/menuitem.048b5a85ccf2cf40a9be6aff100000f7/?vgnextoid=c0c756a575acd010VgnVCM100000bb030a0aRCRD&i18n.http.lang=es&calendarioPqBopa=true>

³²⁵ BOLETÍN OFICIAL DEL ESTADO. (n.d.). Consultada en enero 3, 2013. <http://www.boe.es/>

miembros puso a disposición de la ciudadanía las decisiones judiciales importantes del Poder Judicial en los siguientes términos:

“Poner a disposición del público, bajo forma electrónica fácilmente accesible, las decisiones importantes de la Justicia la obligación de asegurar la exactitud de los textos el acceso gratuito. Cuando la presentación de los textos publicados ha sido mejorada puede ser oportuno hacer pagar un canon y aplicarse a los editores y distribuidores del sector privado la protección de las personas con respecto al tratamiento automatizado de los datos de carácter personal”³²⁶.

Sin embargo, antes de la citada Recomendación el Consejo General del Poder Judicial en 1997 creó el Centro de Documentación Judicial (CENDOJ), que pretendía principalmente remplazar la publicación en medio físico y a su vez controlar el acceso a las decisiones judiciales con relación al manejo otorgado por editoriales con ánimo lucrativo, que en la gran mayoría de casos restringían el acceso de los ciudadanos a la información.

2.15.6. Unión Europea

Dentro de la Unión Europea, hicimos referencia que en las bases de datos jurídicas existentes en el pasado, encontrábamos desde el año 1966 a CELEX, no obstante, en el año de 1981 se convirtió en EUR-Lex, que en la actualidad se encuentra vigente y

³²⁶ PODER JUDICIAL DE ESPAÑA. (n.d.). Consultada en junio 22,2015. Poder Judicial de España : <http://webcache.googleusercontent.com/search?q=cache:543tGJL57A0J:www.poderjudicial.es/stfls/PODERJUDICIAL/DOCTRINA/FICHERO/ACCESIBILIDAD%2520Y%2520DIFUSION%2520DE%2520LA%2520JURISPRUDENCIA%2520EN%2520EL%2520MARCO%2520DE%2520INTERNET%25201.pdf+&cd=2&hl=es&ct=clnk&gl=co>

actualizado hasta la fecha y se puede consultar en su página web, desde el año de 1998, esta base de datos es dirigida por la oficina de publicaciones de la Unión Europea, que es la responsable de presentar sus versiones³²⁷.

CELEX, que era considerada como la base de datos más importante para los usuarios que requerían acceso a la regulación de la Unión Europea, continúa como la abanderada en el tema, con su versión actual, EUR-Lex. No en vano, lleva 46 años desde su creación, esto aunado con la cantidad de documentos recopilados, la condujo a estar considerada como una de las bases de datos de mayor volumen del mundo, toda vez que cada uno de los documentos e información, que en un principio se encontraba disponible para el suscriptor o usuario en once de las veintitrés lenguas oficiales de la UE, hoy en día se publica en todos los idiomas de la Unión europea, además de lo anterior, los documentos jurídicos vienen acompañados por información complementaria que ofrece la posibilidad de comprender de forma amplia el derecho comunitario.

Otro cambio, que surgió de la mano con EUR-Lex, es la posibilidad de acceso gratuito y sin restricción alguna de la ubicación del usuario, toda vez que según la máxima del derecho que establece que "La ignorancia de la ley no es excusa", como respuesta a lo anterior y con ocasión de la celebración de los 25 años de la legislación europea en línea, el 23 de noviembre en la Oficina de Publicaciones Oficiales de la Comunidad Europea, da paso al surgimiento de EUR-Lex, que hoy en día es considerada como la página web más completa en ofrecer acceso a todas las personas.

³²⁷ EUR-LEX. (n.d.). Consultada en marzo 14, 2015.
<http://eur-lex.europa.eu/homepage.html?locale=es>

Dentro de EUR-Lex, se encuentran contenidos los tratados constitutivos de la Unión Europea, la totalidad de los Diarios Oficiales actualizados desde el 1 de enero de 1998 hasta la fecha, la legislación comunitaria vigente, que se realiza con la utilización de los contenidos de CELEX, no obstante, se presentan problemas cuando los formatos de la ley introducen gráficos o tablas. De igual forma, se encuentran incluidos los actos legislativos introductorios de reformas posteriores, los actos jurídicos preparatorios desde el año 2000 y la jurisprudencia de la UE desde el año 1998.

EUR-Lex cuenta con cerca de 2.815.000 documentos, que se actualizan a diario, por lo cual se ha estimado que cada anualidad se incluye aproximadamente 12.000 documentos nuevos³²⁸.

Dentro de EUR-Lex, como productos finales de acceso público contamos con la posibilidad de encontrar los tratados constitutivos (versiones originales y actualizaciones posteriores), “los tratados que los modifican, los tratados de adhesión de cada una de las seis ampliaciones y algunos otros textos esenciales”³²⁹.

Al mismo tiempo, EUR-Lex ofrece la consulta libre de la totalidad de los acuerdos internacionales, que celebran las Comunidades Europeas con ocasión de las funciones y competencias que los facultan para celebrar con Estados miembros en virtud de la competencia compartida, a estos se les denomina acuerdos mixtos, a su vez contiene los acuerdos celebrados con terceros Estados u organizaciones

³²⁸ *Ibíd.*

³²⁹ *Ibíd.*

internacionales, las decisiones de los comités mixtos, creados para la aplicación y gestión de un acuerdo internacional, enlaces a la base de datos de acuerdos del Consejo de la Unión Europea y a la base de datos de la Oficina de Tratados de la Comisión Europea.

En igual sentido, las bases de datos de EUR-Lex, ofrecen al usuario la posibilidad de consultar la legislación vigente y consolidada, así como los trabajos preparatorios, la jurisprudencia de la Unión Europea, jurisprudencia nacional, el diario oficial y las preguntas parlamentarias. Estas últimas, son interrogantes que plantean y dirigen los parlamentarios europeos a la Comisión y al Consejo, preguntas estas que una vez planteadas se publican en el Diario Oficial, con la respuesta de la institución a la que van dirigidas.

Es importante mencionar, que dentro de la página de EUR-Lex, se ofrece la posibilidad de consultar la legislación vigente y consolidada, los acuerdos internacionales y los trabajos preparatorios, con el uso de los ya estudiados tesauros, que gracias a su estructura ofrecen una comprensión global y organizada de la información contenida en la base de datos.

De otra parte, las opciones de búsqueda que ofrece esta página de bancos de datos jurídicos especializados, contiene diferentes opciones de visualización, bien sea de búsqueda simple o búsqueda avanzada, a su vez ofrece la posibilidad de obtener el documento en descarga o para lectura en formato PDF, HTML, DOC y TIFF, según sea la elección del usuario y la disponibilidad en cada caso del documento y presenta a disposición del usuario metadatos analíticos relativos a cada documento.

En vista de la necesidad de informar al ciudadano y ser transparentes, en el año 2001, la Unión Europea elaboró un documento marco contentivo de las políticas de información y comunicación, a la vez que presentó la página web, de acceso libre desde cualquier ordenador del mundo, denominada EUR-lex, en la cual reposan la base de datos jurídica de la legislación y demás documentos jurídicos que regulan a la UE.

Ahora bien, hay quienes afirman, que los *“avances de la UE en este área están directamente relacionados con la integración de Suecia y Finlandia como países miembros en 1998. Los países escandinavos cuentan con una larga tradición administrativa de medidas de acercamiento al ciudadano a los procesos de toma de decisiones.*

*Este aspecto reviste una especial importancia si se quiere fomentar la confianza en unas instituciones de por sí complejas. Y, sobre todo, por la adopción del Reglamento 1049/2001 relativo al acceso del público a los documentos del Parlamento Europeo, del Consejo y de la Comisión (DOCE L 145, 31.5.2001, p. 43-48) cuya entrada en vigor el pasado mes de diciembre de 2001 ha supuesto una pequeña revolución en la administración comunitaria por conceder derechos de acceso a los ciudadanos a la práctica totalidad de la documentación comunitaria, con excepciones relacionadas con asuntos de defensa y seguridad, relaciones internacionales y política económica, financiera y monetaria”*³³⁰.

³³⁰ ALFONSO MOREIRA, C. G. (2002). Las bases de datos de la UE. Revista Valenciana de Economía y Hacienda. Consultada en mayo 31, 2012. p. 249.
http://portales.gva.es/c_economia/web/rveh/pdfs/n5/materiales5.pdf

La Unión Europea³³¹, Cuenta con bases de datos jurídicos sobre la legislación: Directivas, reglamentos y decisiones en curso de adopción o que han entrado en vigor en el territorio europeo.

De igual forma, cada institución europea cuenta con un portal de internet de acceso gratuito, es así como el parlamento europeo en su sitio cuenta con un hipervínculo que ofrece una base de datos jurídica sobre sus archivos³³², en esta página es posible consultar los procedimientos legislativos ordinarios, de aprobación, los procedimientos no legislativos, de cooperación, presupuestarias, de aviso de conformidad, los procedimientos institucionales sobre la nominación o la moción de censura, entre otros.

Así mismo, a nivel presidencial, el Consejo europeo cuenta con su sitio de internet, dentro del cual es posible tener acceso a las reuniones ordinarias y extraordinarias del Consejo Europeo³³³. A nivel ministerial se encuentra el Consejo de la Unión Europea, que contiene un registro de todos los documentos del Consejo desde el 1 de enero de 1999 por un sistema de archivo automático³³⁴. Así, todo documento no sensible sometido al Consejo o a una de sus instancias preparatorias, que pueda servir de base a las deliberaciones, como puede tener una influencia sobre el proceso de la decisión o reflejar los progresos logrados sobre un asunto, es automáticamente mencionado en el registro.

³³¹ L'UNION EUROPÉENNE. (n.d.). Consultada en marzo 15, 2015. L'Union Européenne: http://europa.eu/index_fr.htm

³³² PARLEMENT EUROPÉEN. (2014, enero 18). Archives: <http://www.europarl.europa.eu/parlArchives/comArch/comDocSearch.do;jsessionid=B91E98AD4078138EE6655FB2382A475E.node1?language=fr>

³³³ CONSEJO DE LA UNIÓN EUROPEA. (2014, enero 18). . <http://www.consilium.europa.eu/es/council-eu/>

³³⁴ *Ibíd.*

De otra parte, dentro de las instituciones europeas, se encuentra el sitio de la Comisión Europea, en este sitio es posible tener acceso a la legislación adoptada y la que se encuentra en curso de elaboración³³⁵. En lo relacionado con la jurisprudencia europea, es la Corte de Justicia de la Unión europea la que ofrece un compilado de las decisiones del Tribunal de la función pública, los tribunales y de la Corte de Justicia de la Unión Europea en su sitio de internet³³⁶.

2.15.7. Alemania

De los bancos jurídicos del pasado en Alemania encontramos que el sistema de información jurídica, para la República Federal Alemana, juris continua vigente y constituye una de las recopilaciones de jurisprudencia más extensa en el campo del derecho. La información en juris se organiza de forma individual, según la necesidad, con bases de datos existentes, y presentan servicios para el sector público y soluciones de negocios para firmas de abogados y empresas³³⁷.

En juris la búsqueda de la información se puede realizar mediante el uso de abreviatura de una ley, por ejemplo, Código Penal, Código Civil, etc., los términos del título o del texto legal. Igualmente se puede realizar la búsqueda en los sitios de la Gaceta de Leyes Federales o el número de sitios a prueba A de la Federal Law Journal³³⁸.

³³⁵ COMMISSION EUROPÉENNE. (2014, enero 18). http://ec.europa.eu/legislation/index_fr.htm

³³⁶ TRIBUNAL DE JUSTICIA DE LA UNIÓN EUROPEA. (2014, enero 18). Consultada en agosto 29, 2014. http://curia.europa.eu/jcms/jcms/j_6/

³³⁷ JURIS. (n.d.). Consultada en septiembre 1, 2014. Juris: <http://www.juris.de/jportal/nav/kostenfreieinhalte/infokostenfreieinhalte.jsp>

³³⁸ Ibídem.

Es importante señalar que el acceso a información del marco normativo de Alemania se obtiene mediante el sitio web, página web del Tribunal de Justicia de la Unión Europea³³⁹.

En Alemania encontramos el sitio web German Law, el cual contiene la ley alemana en lengua inglesa. La información la presentan en categorías por sentencias, estatutos, entre otros y cada una de estas pueden ser consultadas por medio de enlaces o hipervínculos que facilitan la consulta de la información. Los enlaces se definen por palabras dentro del texto de la página web, los cuales permiten la dirección a la información requerida y cada uno de estos enlaces apuntan a datos que se encuentran dentro del mismo servidor y dominio en el que se localiza el home del portal³⁴⁰.

De igual forma se encuentra un enlace que direcciona al usuario a la base de datos del servidor, la cual permite buscar información de la jurisprudencia alemana. La información se puede consultar por autor, título, clasificación y nombre de la corte³⁴¹. El acceso a la información es gratuito y la última actualización se hizo en noviembre de 2011, lo que hace pensar que la información registrada se encuentra desactualizada.

También encontramos en Alemania el sitio web Jura, el cual brinda varios enlaces permitiendo el direccionamiento a distintos cuerpos normativos y jurisprudenciales

³³⁹ ASSOCIATION OF THE COUNCILS OF STATE AND SUPREME ADMINISTRATIVE JURISDICTIONS. ACA-Europe. (2015). Consultada en agosto 29, 2014. Bases de Datos: <http://www.aca-europe.eu/index.php/en/dec-nat-en>

³⁴⁰ THE GERMAN LAW ARCHIVE. (2006). Consultada en agosto 29, 2014. <http://www.iuscomp.org/gla>

³⁴¹ *Ibíd.*

de Alemania. Al ingresar al link de la información Alemana, encontramos que en el nivel de las normas se halla una lista de enlaces a los portales de los estados, con las leyes de cada uno de ellos y enlaces a las páginas de la ley federal. A nivel de jurisprudencia, la ley contiene una lista de enlaces a las páginas web de los tribunales de Alemania³⁴².

Es importante señalar que el Ministerio Federal de justicia está desarrollando un proyecto conjunto con juris GmbH, para que los ciudadanos alemanes puedan acceder de forma gratuita, a la legislación federal. En la página web se puede, se publica la información de dicho proyecto y en la cual ya se pueden adelantar algunas búsquedas. En este proyecto la búsqueda de la información se puede hacer haciendo uso de los signos conectivos u operadores lógicos³⁴³.

Otro sitio web, que ofrece más de cien leyes federales es el sitio **dejure.org**, que presenta un listado de enlaces con relación a temas legales, listado que se encuentra jerarquizado en orden alfabético de acuerdo con el tema de cada una de las normas. Igualmente presenta un banner que permite seleccionar las búsquedas con la letra inicial de cada tema objeto de análisis, en el que se registra noticias y actualización de normas³⁴⁴.

En el sitio Bundesgesetzblatt, se encuentra en línea El Boletín Oficial Federal, parte I (desde 1949) y parte II (desde 1951). La búsqueda de la información se puede

³⁴² JURISTISCHES INTERNETPROJEKT SAARBRUCKEN. (n.d.). Consultada en agosto 29, 2014.
<http://www.jura.uni-saarland.de/deutschland/>

³⁴³ BUNDESMINISTERIUM DER JUSTIZ UND FÜR VERBRAUCHERSCHUTZ. (n.d.). Consultada en agosto 30, 2014.
<http://www.gesetze-im-internet.de/titelsuche.html>

³⁴⁴ DER JURE.ORG. Rechtsinformationssysteme GmbH. (n.d.). Consultada en agosto 30, 2014.
<http://dejure.org/>

realizar mediante título, abreviatura, fecha de emisión, año de entrega, entre otros. Adicionalmente, permite la búsqueda del texto completo del documento³⁴⁵.

Se evidencia que en Alemania no existe un único portal que se encargue de centralizar la información relativa a la normatividad, por el contrario, existe una gran variedad de portales y enlaces que ofrecen dicha información, la cual varía en sus métodos de búsqueda y en la manera de organización de los datos.

2.15.8. Venezuela

Los bancos de datos que en Venezuela existieron en el pasado de la informática Jurídica documental, sistema juribanc y los trabajos desarrollados de investigación y aplicación de la informática jurídica de la Asociación Latinoamericana de Informática Jurídica y Derecho Cibernético, en el presente no se encuentran en internet. Se han desarrollado otros como el Sistema de Cooperación Jurídica (SCJ), que vio la necesidad de establecer una mayor comunicación y fortalecer las relaciones entre las unidades de información en el campo jurídico venezolano, y es así como invita a la Biblioteca de la Academia de Ciencias Políticas y Sociales para que realicen un encuentro de bibliotecas y Unidades de Información Jurídica de Caracas, el 16 de octubre de 2003³⁴⁶.

La Biblioteca de la Academia de Ciencias Políticas y Sociales planeó la creación del Sistema de Cooperación Jurídica cuyo objetivo era la integración de las bases de

³⁴⁵ BUNDESANZEIGER VERLAG GMBH. (n.d.). Consultada en agosto 30, 2014. <http://www.bundesanzeiger-verlag.de/ueber-uns.html>

³⁴⁶ SISTEMA DE COOPERACIÓN JURÍDICA. SCJ. (n.d.). Consultada en mayo 29, 2014. <http://www.scjuridica.org.ve/portal/php/001.php>

datos de las bibliotecas jurídicas mediante un portal en la web que diera acceso a un catálogo colectivo en línea, y es así, como se puede consultar los diferentes bancos de datos, mediante un motor de búsqueda amistoso, y en forma simultánea por conducto de un menú único en la web³⁴⁷.

Cumpliendo con su objetivo el Sistema de Cooperación Jurídica, desde octubre de 2004, ha incorporado las bases de datos de la biblioteca de ciencias jurídicas y políticas de la Universidad Central de Venezuela y de la biblioteca Tulio Chiossone del Instituto de Ciencias Penales de la misma universidad. A partir del 2005 cuenta con la participación de la biblioteca Central de la Universidad Católica Andrés Bello³⁴⁸.

El sistema de cooperación jurídica permite realizar búsquedas simultáneas en las diferentes bases de datos de las bibliotecas incorporadas al sistema, garantiza la disponibilidad de una información completa, sistemática y actualizada en el área jurídica nacional y tiene la información disponible las 24 horas en internet³⁴⁹.

Encontramos también en Venezuela el Centro de Investigaciones Jurídicas de la Academia de Ciencias Políticas, quien elaboró y mantiene una base de datos bibliográfica referencial, conocida con el nombre Ulpiano, que contiene tanto la información de publicaciones periódicas jurídicas de Venezuela (revistas, boletines y anuarios), como los artículos de cada una de estas publicaciones, que han sido

³⁴⁷ *Ibídem.*

³⁴⁸ *Ibídem.*

³⁴⁹ *Ibídem.*

clasificados por palabra clave, autor, materia, año y título de la publicación desde el año de 1980³⁵⁰.

Ulpiano contiene referencias de más de diez mil artículos y ciento sesenta y tres publicaciones, tanto activas como discontinuadas. Algunos artículos tienen texto completo, otros un resumen y otros simplemente los datos bibliográficos, pero en todo caso permite encontrar documentos jurídicos y posteriormente con la información que se encuentre ubicar los documentos físicamente o incluso vía web³⁵¹.

Tanto para los artículos como para las publicaciones permite la búsqueda avanzada, en el primer caso se puede hacer la consulta por autor personal, autor institucional, título revista, título artículo, materia, descriptores, editor o editorial y en el segundo caso por título, materia, editor, editorial o lugar de publicación. En el caso de búsqueda de artículos permite seleccionar algún descriptor para limitar la búsqueda. Así mismo permite ir seleccionando artículos y ver al final la lista completa de lo seleccionado³⁵².

En relación con el banco de datos jurídico ULPIANO consideramos que por ser una base de datos referencial no le presta gran utilidad al usuario, puesto que lo obliga a ir a otras fuentes a encontrar el documento completo.

³⁵⁰ ULPIANO. *Primera fuente de referencia jurídica de Venezuela*. (n.d.). Consultada en junio 5, 2014. <http://www.ulpiano.org.ve/>

³⁵¹ *Ibíd.*

³⁵² *Ibíd.*

El Tribunal Supremo de Justicia tiene en su sitio web una base de datos que permite buscar la jurisprudencia y todos los documentos jurídicos relacionados con el Tribunal. La Jurisprudencia y las decisiones, se encuentran en la opción de búsqueda señaladas con su nombre, en donde se halla el registro de todas las providencias y decisiones del Tribunal Supremo de Justicia desde el año 2000 y su búsqueda se puede realizar por fecha, índice temático y criterios múltiples. También es posible la búsqueda mediante las salas que integran el Tribunal y en una fecha determinada³⁵³.

Adicionalmente, es posible consultar tanto las audiencias orales presentadas por las salas Constitucional, Penal, Social, Agraria y Político Administrativa como las cuentas, casos en línea y la legislación³⁵⁴.

En la base de datos jurídica del Tribunal Supremo de Justicia se puede hacer uso de conectores lógicos, siendo necesario escribirlos en el campo de búsqueda, es decir, no facilita la búsqueda mediante el uso de opciones automáticas para el empleo de conectores.

Otra base de datos en línea que encontramos en Venezuela es el catálogo de la biblioteca Andrés Aguilar Mawdsley de la academia de ciencias políticas y sociales, que contiene referencias bibliográficas de libros, folletos, separatas, discursos, tesis y artículos de diferentes publicaciones periódicas de Venezuela e internacionales. En algunos casos, además de los datos bibliográficos también está disponible un resumen o el texto completo del título encontrado. La información se puede obtener

³⁵³ VENEZUELA. TRIBUNAL SUPREMO DE JUSTICIA. (2014). Consultada en junio 5, 2014.
<http://www.tsj.gob.ve/>

³⁵⁴ *Ibíd.*

mediante búsqueda libre, que es la búsqueda por defecto, la cual permite buscar por autor personal, autor institucional, título, tema, país y editor etc., y por búsqueda avanzada, utilizando los signos conectivos “O” para recuperar la información por conducto de disyunciones o “Y” para conjunciones en los campos autor personal, autor institucional, título y temas, entre otros³⁵⁵.

2.15.9. Colombia

Las bases de datos jurídicas que se encuentra en Colombia, son entre otras, Portal Jurídico DMS, esta página web contiene bases de datos jurídicas colombianas y forma parte de la editorial *Ediciones E Investigaciones Ltda*, que dentro de su página de acceso restringido con previa suscripción, ofrece la posibilidad de emprender consultas en los diferentes Códigos, Regímenes, Estatutos, Leyes desde el año 1620 al año 2012, Decretos desde el año 1980 a 2012, jurisprudencia de las altas cortes, minutas y modelos, normativa sobre el sistema penal acusatorio, videos de audiencias reales, doctrina de la Dirección de Impuestos y Aduanas Nacionales DIAN, doctrina de las Superintendencias, legislación internacional, tratados internacionales, normativa de la ciudad de Bogotá, compilación de normas de derecho disciplinario y un diccionario jurídico³⁵⁶.

³⁵⁵ BIBLIOTECA ANDRÉS AGUILAR MAWDSLEY. (2011). Consultada en junio 23, 2014.
<http://acienpol.msinfo.info/>

³⁵⁶ PORTAL JURÍDICO DMS. (n.d.). <http://www.portaldms.com/index.html>

También encontramos a Lexbase S.A.S., sitio web que se creó en 1992. Este sitio web ofrece la consulta de Boletines, Normas, Jurisprudencia, Códigos Estatutos, Normatividad Regional, Doctrina, entre otros³⁵⁷.

Así mismo, hallamos a Leyex.info, organizado por módulos de consulta y ofrece en su banco de datos Leyes, Decretos, Códigos, Circulares, Resoluciones, Jurisprudencia, Conpes, Laudos Arbitrales, Proyectos de Ley, Comunicados de Prensa de la Corte Constitucional, Tratados internacionales, Diario Oficial, Normativa internacional, Derecho Comparado, Noticias Económicas y Jurídicas, Revistas Económicas y Jurídicas y la Constitución Política de Colombia. La información mencionada la presentan clasificada, concordada, organizada y actualizada³⁵⁸.

Otro sitio de información jurídica importante en Colombia es notinet.co, consiste en un noticiero jurídico normativo, vía Internet. Inicio actividades desde 1998, procesando información jurídica proveniente de las altas cortes y de los distintos entes gubernamentales, en materias como administración pública, comercio exterior, servicios públicos, derecho penal, laboral, financiero y tributario³⁵⁹.

En Colombia también hallamos la base de datos de multilegis, que ofrece acceso a una colección de códigos básicos, comercio exterior, derecho financiero, laboral,

³⁵⁷ LEXBASE. (n.d.). Consultada en agosto 28, 2014. <http://www.lexbase.com/>

³⁵⁸ LEYES.INFO. Información Económica y Jurídica de Colombia. (2015). Consultada en agosto 28, 2014. http://www.leyex.info/index_descripcion.php

³⁵⁹ NOTI.NET. (2015). Consultada en agosto 28, 2014. <http://www.notinet.com.co/>

privado, público tributario y contable. Incluye la colección de legislación y jurisprudencia colombiana, minutas, modelos, entre otras obras³⁶⁰.

La base de datos de Ambientalex.info, proporciona información ambiental nacional e internacional. Brinda información sobre Normas Nacionales, Códigos, Jurisprudencia, Derecho Internacional Público, Normas Regionales, Información Científica y Técnica, Guías Ambientales, Mapas SIGAMB, Noticias Ambientales y Jurídicas, entre otros³⁶¹. Esta base de datos es actualizada vía internet, diariamente.

Asimismo, encontramos la Imprenta Nacional, que una de sus funciones es la publicación oportuna de la gaceta del Congreso, en donde se reseña no sólo el origen y desarrollo de las leyes del país, sino también los debates parlamentarios. Permite la consulta tanto del último diario oficial como el histórico de los mismos. La búsqueda de la información se puede realizar por número, rango de fechas, norma y palabras y la información la recupera en texto completo³⁶².

También tenemos el sitio web todo el derecho.com, cuya base de datos brinda acceso a la Constitución Política, decretos de conmoción interior, estatutos y códigos nacionales, legislación por materia, por orden cronológico y por tipo de norma, normas de comercio exterior, derecho financiero, laboral, privado, público tributario

³⁶⁰ MULTILEGIS. CONEXIÓN A LA INFORMACIÓN JURÍDICA. (2011). Consultada en agosto 28, 2014.
<http://www.multilegis.com/colombia.html>

³⁶¹ AMBIENTALEX.INFO. El portal ambiental. (2015). Consultada en agosto 28, 2014.
<http://www.ambientalex.info/>

³⁶² IMPRENTA NACIONAL DE COLOMBIA. (2015). Consultada en agosto 30, 2014.
<http://www.imprenta.gov.co/portal/page/portal/inicio>

y contable. Adicionalmente incluye una colección de jurisprudencia, minutas, modelos, entre otras obras³⁶³.

No sobra indicar que los sitios web de la Corte Suprema de Justicia, Consejo de Estado y Corte Constitucional suministran bases de datos de la jurisprudencia emitida por ellos. Cabe mencionar, que a pesar de la existencia de esta innumerable cantidad de bancos de datos, son pocos aquellos que se han dado a la tarea de ampliar los criterios de búsqueda, por lo cual, es necesario mencionar que en bases de datos como la de LEGIS resulta difícil emprender búsquedas con el desconocimiento del número exacto de la norma, esto sin obviar que cuando se emprenden búsquedas por conducto de palabras clave, estas en la gran mayoría de las ocasiones resultan infructuosas, toda vez que se debe conocer con exactitud los descriptores impresos en las bases de datos.

Ahora bien, como se puede examinar en todas y cada una de las bases de datos en las páginas web citadas con anterioridad, los contenidos no van más allá de la incorporación del documento jurídico tal y como lo profiere la entidad u órgano correspondiente, por lo cual es casi imposible encontrar información que además cuente con un análisis jurídico de fondo, lo anterior, para los efectos de optimización e implementación de herramientas semánticas resulta indispensable tal y como lo veremos en el capítulo tercero de esta tesis.

³⁶³ TODOELDERECHO.COM. (2014). Consultada en julio 6, 2014. <http://bases.todoelderecho.com>

2.15.10. Dinamarca

“En Dinamarca, Lex Dania es el sistema nacional para la creación y el intercambio de documentación legislativa. Se trata de un sistema sustentado en el modelo OIO-XML (<http://isb.oio.dk/info>) que dio sus primeros pasos en el Ministerio de Ciencia Tecnología e Innovación y fue adoptado por el Ministerio de Justicia y el Parlamento danés”³⁶⁴.

2.15.11. Países Bajos

Desde la Universidad de Amsterdam se creó en 2002 el portal METALEX (<http://www.metalex.eu>), en la actualidad un servidor de documentos y herramientas en formato abierto que sirve de marco genérico y extensible para la codificación de la estructura y contenidos de documentos legales en XML y RDF³⁶⁵. METALEX es igualmente el origen del proyecto europeo ESTRELLA: European project for Standardized Transparent Representations in order to Extend Legal Accessibility³⁶⁶

En este periplo es obligado reconocer un punto de encuentro trascendental para difundir proyectos e iniciativas, los Legislative XMLWorkshops, realizados en el marco de los encuentros de la Foundation for Legal Knowledge Based Systems³⁶⁷.

³⁶⁴ ALVITE DÍEZ, M. L. (2009). op.cit. p. 39.

³⁶⁵ CEN METALEX. Open XML Interchange Format for Legal and Legislative Resources. (2006, Julio 7). Consultada en junio 22, 2015. <http://www.metalex.eu/>.

³⁶⁶ THE ESTRELLA PROJECT WEBSITE. (n.d.). Consultada en junio 22, 2015. <http://www.estrellaproject.org/>

³⁶⁷ JURIX. THE FOUNDATION FOR LEGAL KNOWLEDGE BASED SYSTEMS. . (2015). Consultada en junio 22, 2015. <http://jurix.nl/>

ODS: Official Documents of the United Nations: Base de datos en línea, que reúne los documentos para reuniones de las Naciones Unidas publicados desde 1946. Ofrece cuatro sistemas de búsqueda a saber: Búsqueda simple, búsqueda avanzada, búsqueda básica y búsqueda global³⁶⁸.

El SAD abarca todos los tipos de documentos oficiales de las Naciones Unidas publicados a partir de 1993. Sin embargo, se están agregando a diario documentos publicados anteriormente. El SAD también permite consultar las resoluciones de la Asamblea General, el Consejo de Seguridad, el Consejo Económico y Social y el Consejo de Administración Fiduciaria aprobadas a partir de 1946. El sistema no incluye los comunicados de prensa, las publicaciones de las Naciones Unidas para la venta, la United Nations Treaty Series ni los folletos del Departamento de Información Pública. Una donación japonesa de más de 30.000 documentos digitalizados ha sido añadida al SAD.

El SAD es un sistema multilingüe, por lo tanto, exige navegadores de Internet que se ajusten a la norma UNICODE, como MS Internet Explorer 5.5, Netscape 6.21 o versiones posteriores. Para consultar el SAD, es necesario asegurarse que el navegador cumpla con ese requisito.

2.15.12. Estados Unidos

En los Estados Unidos encontramos el Sistema LEXIS, que su origen se remonta al año 1964, como lo expresa Bing, fruto de una iniciativa de la *Ohio Bar Association*

³⁶⁸ SISTEMA DE ARCHIVO DE DOCUMENTOS DE LAS NACIONES UNIDAS. (2004). Consultada en junio 22, 2015. <http://documents.un.org/>

quien contrata a *Data Corporation* de Dayton (Ohio) para desarrollar el sistema de recuperación. En 1968 fue adquirida por *Mead Corporation*, naciendo *Mead Data Central* que lanzó un prototipo del sistema operativo al año siguiente. El nuevo servicio apareció en 1973 bajo el nombre de LEXIS, preocupado de modo especial por la documentación judicial y por la creación de un sistema de recuperación fiable y robusta que, incluso es “creado por la Mead Data Central en 1963 Es el banco de Datos Jurídico más grande del mundo, con casi 30 billones de caracteres cargados y con una entrada anual en memoria de alrededor de 35 mil sentencias memorizadas en texto integral. Sistema de carácter privado”³⁶⁹.

En la actualidad,

“mantiene muchas semejanzas con la versión inicial. Su principal competidor se desarrolló en el seno de la editorial jurídica más importante de Estados Unidos, West Publishing que sacó a la luz en 1975 su sistema WESTLAW. El software de recuperación inicial, sobre el que se realizaron desarrollos profundos, fue adquirido a la empresa canadiense QL Systems. En Canadá, el desarrollo guarda relación con el proceso de Estados Unidos en virtud del uso inicial del software QL System, originariamente denominado QUIC/LAW que nace en 1968 de la mano del profesor Hugh Lawford. El aspecto más singular del sistema residía en que la recuperación se basaba en algoritmos de ordenación que atendían a la frecuencia de aparición de las palabras. En 1974 se prefirió dar prioridad a los conectores

³⁶⁹ VARGAS OSORNO T.G. ob. cit. p. 86.

booleanos y a los operadores de distancia, reduciendo a simples opciones los algoritmos de ordenación”³⁷⁰.

Lexis es un sistema interactivo de texto completo. Como el sistema es interactivo el abogado puede mantener un diálogo continuo con el computador, por medio de un terminal³⁷¹.

El sistema Lexis ofrece la colección más amplia de contenidos en línea, con textos integrales y actualmente no sólo opera en los Estados Unidos, sino también en Inglaterra y Francia, países en donde se han almacenado las decisiones del Consejo de Estado³⁷².

Se creó el sistema Lexis Nexis, universo académico, (Lexis Academic Universe) mediante el cual se presta el servicio de información norteamericana en línea amplia y completa. Este servicio ofrece acceso a la información jurídica norteamericana permitiendo búsquedas de texto integral del material jurídico (Legal research), con cobertura federal, estatal y en algunos casos internacional³⁷³.

Adicionalmente, se encuentra el **sistema iuris**, que comienza en 1970 cuando el Departamento de Justicia inicia investigaciones y comienza a desarrollar esfuerzos

³⁷⁰ ALVITE DÍEZ, M. L. (2003). op.cit. p. 192.

³⁷¹ VARGAS OSORNO T.G. ob. cit. p. 86.

³⁷² *Informática Jurídica Documentaria*. (n.d.). In [Informática instrumento del derecho]. Consultada en julio 2, 2013. <http://biblio.juridicas.unam.mx/libros/4/1941/7.pdf>.

³⁷³ LEXISNEXIS ACADEMIC. (2015). Consultada en junio 22, 2015. LexisNexis Academic: <http://www.lexisnexis.com/en-us/products/lexisnexis-academic.page>.

tendientes a la preparación de un sistema automatizado para recuperar información jurídica. Sistema de carácter oficial³⁷⁴.

Iuris es un sistema que hace posible la recuperación del texto completo (*full-text*) de los documentos necesarios para una investigación jurídica, a partir de una fuente central de información. Se trata de un sistema *on line*, es decir, la información puede ser recuperada del archivo de documentos por medio de una llamada telefónica³⁷⁵.

También está el Sistema *Flite*, Sistema de carácter oficial. Antiguo *Lite* (Federal Legal Information Through Electronics), desarrollado a mediados de los años 60 por la fuerza aérea. Este sistema opera con una modalidad Batch, esto es, solamente puede ser satisfecha y atendida una solicitud en cada oportunidad. Sus archivos incluyen el *United State Code*, las decisiones de la Corte de Apelaciones de los Estados Unidos, los reportes de la Corte marcial y manual para cortes marciales, tratados y acuerdos internacionales de especial interés para el Departamento de Defensa³⁷⁶.

Actualmente, la colección de *Flite* tiene más de siete mil (7000) opiniones de la Corte Suprema que data desde 1937. Lo anterior, está disponible en línea directa de dos fuentes: la página de decisiones FedWorld/FLITE Corte Suprema, una producción del Departamento de comercio de Estados Unidos y el Villanova Universidad Escuela de ley FLITE, base de datos de archivo. El sitio de Villanova ofrece una función especial que consiste en que cualquier decisión en la base de

³⁷⁴ VARGAS OSORNO T.G. ob. cit. p. 85.

³⁷⁵ *Ibíd.*

³⁷⁶ *Ibíd.* p. 86.

datos de Villanova FLITE pueda ser enviada gratuitamente por fax. La estrategia de búsqueda es muy sencilla, se puede hacer por el nombre del caso de la decisión que se esté buscando o texto completo, en donde hay que tener sumo cuidado, ya que por el tamaño de la base de datos y el número de palabras que corresponden a 14,5 millones hace que se requiera más de una búsqueda para encontrar la información solicitada. Esta base de datos no es posible interrogarla por las palabras vacías, ya que su motor de búsqueda las ignora³⁷⁷.

Las decisiones se encuentran disponibles como archivos de texto ASCII, que pueden leerse en la pantalla del navegador o guardar en su disco duro y se accede usando la mayoría de programas de procesador de palabras³⁷⁸.

En los Estados Unidos también encontramos AUTO-CITE de Lawyer's Cooperative Publishing Company, que desde 1970 brinda a los abogados norteamericanos la posibilidad de obtener en forma automática las citas de jurisprudencia, reconstruidas históricamente de un archivo automatizado que contiene más de cuatro millones de referencias³⁷⁹.

El Centro de Law Research Service, Comienza en 1964. Utiliza *abstract*. El sistema parte de una clasificación por materias para cada una de las cuales se ha elaborado un *thesaurus* que, ordena los diversos materiales memorizados numéricamente. El usuario está conectado a una red de télex³⁸⁰.

³⁷⁷ THE FEDWORLD A PROGRAM OF THE UNITED STATES DEPARTMENT OF COMMERCE. (n.d.). Consultada en julio 19, 2013. Supreme Court Decisions: <http://supcourt.ntis.gov/>

³⁷⁸ *Ibíd.*

³⁷⁹ GORRÓN A., L. M. (2009, Diciembre 01). *Informatica Jurídica Documental*. Consultada en julio 25, 2013. Slide Share: <http://www.slideshare.net/guest9ca8c4/informatica-juridica-documental>

³⁸⁰ *Ibíd.*

El primer servicio de documentación electrónica de textos jurídicos que entró en funcionamiento en los Estados Unidos, Comienza en 1962, con el Antitrust División of the Department of Justice, quien es responsable del cumplimiento de las leyes antimonopolio de los Estados Unidos. “Funciona con un sistema de *abstract* y comprende desde las nociones jurídicas más generales a las más precisas, siempre en materias relacionadas con la jurisprudencia, legislación y doctrina antitrust.”³⁸¹

LawGuru, es uno de los portales importantes con información jurídica de los Estados Unidos. Es una herramienta de investigación legal sin costo, cuenta con recursos para cada estado, incluyendo casos, opiniones de la corte, códigos, estatutos, facturas y la consulta se hace mediante una lista ordenada alfabéticamente. Tiene un recurso adicional que consiste en la posibilidad de búsqueda legal de preguntas y respuestas en más de 45 áreas del derecho³⁸².

En este portal también se encuentra la biblioteca de derecho de internet (antes los Estados Unidos cámara de representantes de Internet ley biblioteca), que fue ofrecida originalmente al público por cortesía de la oficina del Consejo de revisión de ley de la cámara de representantes de Estados Unidos y que su cometido consistía en que la ley estuviera disponible al público. El sitio proporciona enlaces a un gran número de recursos legales relativos a América (federal y estatal) y leyes extranjeras³⁸³.

³⁸¹ *Ibíd.*

³⁸² LAWGURU. (1997). Consultada en agosto 2, 2013. LawGuru: <http://www.lawguru.com/>

³⁸³ *Ibíd.*

El Tribunal Supremo de Estados Unidos tiene una variedad de opciones en línea disponibles, es así como FindLaw, Inc ofrece una base de datos de las decisiones de esta Corte que se remonta desde 1890 hasta el presente, cuyo servicio es gratuito y permite la consulta por año, número del volumen de U. S. informes y citación, caso título y texto completo.

Las decisiones de la Corte Suprema a partir de 1990 se encuentran a disposición del público, bajo la tutela del proyecto Hermes. En el periodo anterior de 1990 estaban disponibles en una variedad de fuentes on-Net, en distintos formatos y es así como se halla la colección LII de decisiones históricas de la Corte Suprema de los EE.UU. que contiene más de 300 de las decisiones más importantes a través de todo el período de su existencia, se pueden utilizar en formato de cd-rom. También vemos a FedWorld que proporciona punteros a distintos usos de la base de datos FLITE. FLITE cubre el período de 1937 a 1975 y lo hace de forma global. El servicio de USSC + de Infosynthesis proporciona cobertura total a partir de 1966 en adelante y unos 450 casos más antiguos que data de 1793. Por último encontramos a WESTDOC, servicio que ofrece una cobertura completa de las decisiones de la Corte³⁸⁴.

Findlaw, es una empresa de Thomson Reuters, líder en el ofrecimiento de información jurídica en línea y servicios de abogados, empresas y particulares. Las raíces de FindLaw aparecen en 1995, cuando dos abogados compilaron una lista de los recursos de Internet para un grupo de bibliotecarios de derecho en el norte de California. La respuesta del material fue tan positiva que decidieron publicar la

³⁸⁴ CORNELL LAW SCHOOL. (1992). Consultada en septiembre 2, 2013. <https://www.law.cornell.edu/>

información en Internet, y es así como el 9 de enero de 1996, FindLaw.com ofrece mezcla de casos, estatutos, noticias legales, un directorio de abogados, un centro de carreras en línea y herramientas orientadas a la comunidad, tales como las listas de correo y foros de mensajes. El sitio Web se desarrolló rápidamente y se convirtió en uno de los principales sitios de información jurídica en internet, siendo uno de los sitios más populares en donde se puede encontrar información legal gratis, con más de cinco millones de visitas mensuales y el directorio legal más grande disponible.

Findlaw cuenta además con sitios galardonados Web, vídeo en línea, perfiles de abogado en línea, marketing en buscadores y publicidad en Internet³⁸⁵.

Para los investigadores en internet la Base de datos de la Suprema Corte de FindLaw representa una excelente opción por su capacidad y sus eficaces opciones de búsqueda, ya que se muestran y se cargan las opiniones en su totalidad, lo que hace que sea mucho más simple y factible la búsqueda de la información, la impresión y almacenamiento del texto completo rápidamente³⁸⁶.

El sitio WEB de Findlaw incluye jurisprudencia, leyes estatales y federales, un directorio de abogados, noticias y análisis legal e También incluye una revista jurídica gratuita llamada escritura, cuyos contribuyentes (principalmente académicos legales) discuten, explican y debaten asuntos jurídicos de actualidad.

³⁸⁵ FINDLAW. (2015). Consultada en agosto 30, 2013. <http://public.findlaw.com/>

³⁸⁶ FINDLAW. SUPREMA CORTE. (2015). Consultada en octubre 3, 2013. www.ilrg.com/caselaw

En los Estados Unidos también encontramos empresas como infosynthesis³⁸⁷ que ha publicado un CD-ROM con 8599 textos completos de las decisiones de la Corte Suprema de los Estados Unidos, desde 1938 hasta junio de 1998. Los usuarios de internet pueden descargar un dictamen en formato RTF o en formato de documento de Adobe PDF³⁸⁸.

La disponibilidad de opiniones judiciales en línea en los Estados Unidos ha aumentado sustancialmente en los últimos siete años. La mayoría de los tribunales federales, incluyendo los tribunales del circuito y muchos tribunales estatales, podemos afirmar que ahora hay muchos más que tienen páginas Internet patrocinadas por varias entidades educativas y comerciales; sin embargo, las bases de datos de estas opiniones judiciales tienden sólo a cubrir los últimos años. La excepción al problema de la disponibilidad limitada de jurisprudencia es la Corte Suprema de Estados Unidos, que cuenta con decisiones en línea desde 1937 y se encuentran en múltiples fuentes³⁸⁹.

Recapitulando podemos afirmar que en el sector privado las bases de datos más amplias con acceso al internet de opiniones judiciales a nivel federal y estatal en la actualidad sólo están disponibles a través de proveedores privados basados en honorarios, como Lexis-Nexis, Westlaw y Versuslaw. West Publishing ha lanzado Westlaw con tarjeta de crédito que permite la recuperación de los casos estatales y federales y en el sector público la oficina administrativa de los tribunales de los

³⁸⁷ INFOSYNTHESIS PUTS U.S. SUPREME COURT DECISIONS. (1998, Noviembre). Consultada en Septiembre 2, 2013. <http://connection.ebscohost.com/c/articles/1296630/infosynthesis-puts-u-s-supreme-court-%20decisions-cd>.

³⁸⁸ INTERNET LEGAL RESEARCH GROUP. Consultada en Septiembre 23, 2013. <http://www.ilrg.com/caselaw/>

³⁸⁹ *Ibíd.*

EE.UU. mantiene el Directorio de Servicios Públicos de Acceso Electrónico, que proporciona información sobre el acceso automatizado a los EE.UU. y la información de la corte federal del estado y los registros.

Casi todos los tribunales en los Estados Unidos tiene un BBS de acceso telefónico, pero casi ninguno de estos tribunales ha hecho la transición a Internet. El más importante servicio gubernamental coordinada de las opiniones judiciales es PACER (Acceso Público a la Corte Electronic Records), que permite a cualquier usuario con un ordenador personal y un módem para marcar a un distrito o una computadora corte de bancarrota y recuperar la información del caso electrónico oficial y expedientes judiciales. En la mayoría de los tribunales se ha iniciado que los usuarios de este servicio incurran en los gastos de cuota³⁹⁰.

En los Estados Unidos la Universidad de Cornell Legal Information Institute ha puesto las reglas de la Corte Suprema de EE.UU. en Internet y la empresa ILRG ofrece una colección de perfiles biográficos de los actuales jueces de la Corte Suprema³⁹¹.

Ahora bien, para los Tribunales del Circuito Federal más de un sitio web se encuentra a disposición del público en donde se pueden consultar las opiniones y descargar el texto. Cuentan con un enlace de hipertexto que representa la más completa base de datos actual, de libre acceso a través de Internet³⁹².

³⁹⁰ *Ibíd.*

³⁹¹ *Ibíd.*

³⁹² *Ibíd.*

Entre otros centros que prestan el servicio de información jurídica en internet, encontramos al Centro de Law Research Service, que cuenta con uno de los portales legales de internet más grandes y confiables. Sus bases de datos jurídicas disponen de un motor de búsqueda que permite navegar por cualquier criterio de su elección y admite hacer consultas inteligentes en las que primero revisan la base de datos para identificar los registros reales y luego hacen un resumen de dichos datos para el usuario. Como resultado las formas inteligentes nunca presentan resultados vacíos, siempre muestran resultados significativos³⁹³.

2.15.13. Italia

En Italia como en la mayoría de los países, los bancos de datos jurídicos se encuentran en Internet, es así como el centro de documentación automática, centro existente en los inicios de la informática jurídica hoy existen en la red³⁹⁴.

Igualmente, de los bancos jurídicos del pasado encontramos en la web el centro de documentación electrónica de la Suprema Corte di Cassazione, Italgiure Find, es un sistema de interfaz web basada en la investigación y permite la consulta del banco de datos contentivo de documentos de jurisprudencia, doctrina y ley de la CED de la Corte de Casación.

El acceso se encuentra regulado por el Decreto Presidencial 322 de 1981, modificado por el 759 de 1985 y uno reciente de 7 de febrero de 2006 DML³⁹⁵.

³⁹³ LEGAL RESEARCH SERVICES. (2009). Consultada en octubre 3, 2013.
<http://www.lawresearchservices.com>.

³⁹⁴ PARLAMENTO ITALIANO. CAMERA DEI DIPUTATI. (n.d.). Consultada en enero 20, 2014.
<http://goo.gl/dND3kQ>

Italgire Web³⁹⁶, es un banco de datos de jurisprudencia y legislación elaborado por el Centro Elettronico de Documnetazione. CED., de la Corte de Casación Italiana, al cual se puede acceder por internet. Una parte es de acceso confidencial, puesto que hay que reconocer un pago para ingresar, pero otra parte, es de libre acceso para temas que son de roles institucionales³⁹⁷.

El Sistema Italgire Find, en relación con sus inicios ha mejorado notablemente en el sentido que ahora es más funcional, más fácil y más intuitivo. En el pasado solo era posible su consulta por conducto de intranet, ahora se puede acceder a través de internet, dando la posibilidad de ingresar a ella desde cualquier estación de trabajo que este colgada en la web. Las decisiones judiciales se encuentran en textos completos tanto en formato de texto como formato de imagen, lo que permite que el usuario pueda imprimir una copia de la sentencia del documento original. A pesar del tiempo que se requiere antes de ingresar el documento a la base de datos, para su preparación, clasificación y control, se ha reducido el tiempo entre el momento de publicación de la sentencia y el momento de inserción de la misma al banco de datos. Adicionalmente la base de datos también se puede consultar en inglés y francés³⁹⁸.

En Italia se encuentra la gaceta oficial que desde el 1 de enero de 2013, gracias al Ministerio de Economía y Hacienda que de acuerdo con el Ministerio de Justicia y

³⁹⁵ ITALIA, CORTE DI CASSAZIONE. (n.d.). Consultada en enero 27, 2014.

<http://www.cortedicassazione.it/AreaRiservata/>. BDGiurisprudenza/BDGiurisprudenza.asp.

³⁹⁶ ITALGIUREWEB. (n.d.). Consultada en junio 23, 2015. www.italgiure.giustizia.it

³⁹⁷ MARCO, M. A. (n.d.). L'informatica giuridica ed il sistema Italgire. Consultada en enero 27, 2014, <http://www.corsomagistratitributari.unimi.it/STRUTTURA/DOTTRINA%20E%20RELAZIONI/VARIA/Approfondimenti/40%20Alma/A-40%20Alma.pdf>

³⁹⁸ *Ibidem*.

con el apoyo de la Casa de la Moneda Estatal hacen que, el boletín oficial se pueda consultar gratuitamente en formato digital.

Adicionalmente, han desarrollado un proceso integrado que permite garantizar la autenticidad de la información y que, a partir de 2009, se utilice herramientas avanzadas tales como sello digital, que se coloca en cada página de la publicación y firma digital por el jefe del proceso de publicación, que asegura el cumplimiento del contenido de la versión digital y que garantiza que la suscripción no ha sufrido ninguna modificación de la versión en papel³⁹⁹.

El diario oficial se distribuye electrónicamente por conducto del sitio de la Gazzetta Ufficiale electrónica, es una base de datos de gran importancia que contiene los diarios oficiales en texto y archivo PDF, lo que permite ver las publicaciones individuales o, en el caso de la versión de texto, los actos individuales. Sus búsquedas se hacen a través de palabras significativas. El sistema permite encontrar la documentación oficial publicada en la fecha indicada⁴⁰⁰.

El portal Diritto & Diritti, portal jurídico, es una revista en línea que ofrece información jurídica facilitando las sentencias y leyes italianas. Así mismo brinda un boletín de actualización sobre los cambios más importantes que se han dado en el mundo del derecho⁴⁰¹.

³⁹⁹ ITALIA. GAZZETTA UFFICIALE. (n.d.). Consultada en marzo 6, 2014. <http://www.gazzettaufficiale.it/>

⁴⁰⁰ *Ibidem*.

⁴⁰¹ DIRITTO.IT. (n.d.). Consultada en marzo 15, 2014. www.diritto.it

En el año de 1986 la Universidad de Bolonia creó el Centro Interdepartamental, es uno de los principales centros en el área de las ciencias de la computación y la ley. Promueve la investigación en las áreas de filosofía del derecho, sociología del derecho, historia del derecho, bioética e informática jurídica⁴⁰².

También encontramos en línea el Instituto de Información Legal Teoría y Técnicas (ITTIG), es un órgano del Consejo Nacional de Investigación, que nació en 1968 como un instituto para la documentación (IDG) Y adopta su nombre actual en 2002. ITTIG desarrolla investigaciones en el campo de Informática Jurídica y Derecho de la Informática. Frecuentemente interactúa con el estudio académico y científico y la investigación aplicada sobre la relación entre la ley y la ciencia y de las tecnologías de la información y la comunicación legal. El instituto tiene una relación estrecha de colaboración con el Gobierno en la administración electrónica, crea y distribuye bases de datos jurídicas y software de diseño especializado y herramientas para transmitir la información jurídica en la red y participación de los datos del gobierno⁴⁰³.

En conclusión, el ITTIG investiga y aplica tecnologías de información y comunicación a áreas de la ley, lenguaje legal, técnica legislativa, toma de decisiones jurídicas, formación de abogados, derecho e información pública y la política de la

⁴⁰² UNIVERSITÀ DI BOLOGNA. Centro Interdipartimentale di Ricerca in Storia del Diritto, Filosofia e Sociologia del Diritto e Informatica Giuridica. CIRSFID. (2015). Consultada en marzo 20, 2014. <http://www.cirsfid.unibo.it/>

⁴⁰³ CONSIGLIO NAZIONALE DELLA RECHERCHE. Istituto di Teoria e Tecniche dell'Informazione Giuridica. (n.d.). Consultada en mayo 8, 2014. www.cnr.it/istituti/Descrizione.html?cds=104

computación. El Instituto se encuentra en Florencia, y está dirigido por la Dra. Daniela Tiscornia⁴⁰⁴.

La gestión, actualización y distribución de la ley y las bases de datos de abogados están a disposición del público en línea y a través de CD ROM. Los archivos electrónicos se encuentran en el servidor del ITTIG y este Instituto también ha preparado guías que permite la recuperación de sitios web de interés legal⁴⁰⁵.

Es importante resaltar que el ITTIG tiene proyectos de investigación sobre leyes en línea cuyo objetivo es fomentar la transparencia y acceso a la información jurídica de los ciudadanos comunes y corrientes. Es decir, lo que pretende es crear un sitio en internet que constituya el punto de acceso unificado a toda la documentación legislativa de los organismos institucionales italianas de forma gratuita y sin necesidad de ayuda, con un sistema de recuperación de fácil uso y eficiente. Lo anterior proporciona una herramienta que permite a los ciudadanos ejercer su derecho a la información⁴⁰⁶.

Así mismo encontramos el sitio web Norme in Rete (leyes en línea), que ofrece acceso gratuito a la información legal italiana, cuyo objetivo es fomentar el acceso y transparencia de la información jurídica a todos los ciudadanos y construir un sitio en internet que sirviera como punto de acceso unificado para la consulta de toda la

⁴⁰⁴ ITTIG. INSTITUTE OF LEGAL INFORMATION THEORY AND TECHNIQUES. (n.d.). Consultada en mayo 7, 2014.
<http://www.ittig.cnr.it/>

⁴⁰⁵ *Ibidem*.

⁴⁰⁶ *Ibidem*.

documentación legislativa publicada en los sitios web de los organismos institucionales italianos, de fácil acceso y por lo mismo sin necesidad de ayuda⁴⁰⁷.

2.15.14. Brasil

En este país, continúa vigente el sistema PRODASEM, Centro de Informática y Procesamiento de Datos del Senado Federal, su historia se remonta a finales de 1970, cuando el senador João Cleofas, nombró un grupo de trabajo para rendir un informe sobre la viabilidad del proyecto de desarrollo de un centro de computación que estuviera conectado al Congreso Nacional, con el fin de mejorar la eficiencia y eficacia del proceso legislativo⁴⁰⁸.

ProdaseM, implementó tecnología con competencia y técnicas de tratamiento documental, con una visión futurista, llegando a significar un proyecto estratégico, logrando alcanzar visibilidad y reconocimiento nacional e internacional. Desde el principio, se organizó un equipo multidisciplinario con técnicos en informática y sistemas de información, que unidos a asesores legislativos y varios órganos implementaron un sistema con referencias legislativas, legales y documentación, análisis, catalogación y almacenamiento. Lo anterior, permitió un paso importante para la modernización, y hace coherente y precisa la búsqueda y recuperación de información legislativa y decisiones judiciales del Brasil⁴⁰⁹.

⁴⁰⁷ Ibídem.

⁴⁰⁸ BRASIL. SENADO. PRODASEN. (n.d.). Consultada en junio 24, 2014.
<http://www.senado.gov.br/senado/prodaseM/>

⁴⁰⁹ Ibídem.

Para el 12 de octubre de 1972, fecha en la que fue inaugurado prodasem⁴¹⁰, ya estaban desarrollados los módulos del sistema de información legislativa y el sistema administrativo e integrado de bases de datos NJUR, que contiene las disposiciones legales; MATE, para el procesamiento de materiales y BIBR, que son las referencias bibliográficas de la biblioteca del Senado y otros⁴¹¹.

En el desarrollo de Prodasen, en el año de 1990 comienza un agudo cambio de diseño en la lógica computacional de la institución, una arquitectura concentrada y cimentada en mainframe, y es así como evoluciona a una red distribuida y basada en servidores y PCs. Lo anterior, hace parte de la misión del Prodasen, que consiste en dar soluciones tecnológicas, para lograr la consulta de la información y comunicaciones hacia el Senado, de esta forma contribuye a la modernización en las prácticas del proceso legislativo y las actividades de apoyo administrativo, reduce el tiempo y mejora los procesos de trabajo de los legisladores brasileños⁴¹².

Brasil es miembro de la red-desc, red internacional para los derechos económicos, sociales y culturales; quienes cuentan con una base de datos de jurisprudencia y decisiones relacionadas con dichos derechos⁴¹³.

En Brasil también se encuentra el sistema DJI, índice de la ley fundamental de derecho, que incluye definiciones con legislación y jurisprudencia relacionada. Es un diccionario brasileño en línea⁴¹⁴.

⁴¹⁰ VARGAS OSORNO T.G. ob. cit. p. 90.

⁴¹¹ BRASIL. SENADO. PRODASEN. (n.d.). Consultada en junio 24, 2014. <http://www.senado.gov.br/senado/prodasen/>

⁴¹² Ibídem.

⁴¹³ INTERNATIONAL NETWORK FOR ECONOMIC, SOCIAL & CULTURAL RIGHTS. ESCR-NET. (n.d.). Consultada en julio 5, 2014. <http://www.escr-net.org/>

2.15.15. Argentina

En el desarrollo de la informática jurídica documental, Argentina ha sido pionera en América del Sur. Existen bancos de datos jurídicos que permiten la consulta de documentos mediante tesauros jurídicos, rigurosamente elaborados, con herramientas de consulta avanzadas que permiten al usuario llegar en forma acertada y fácil al documento deseado, pero también existen otros que se limitan a la publicación de documentos jurídicos, permitiendo la consulta sólo por número de referencia o el tema, e incluso algunos solo generan enlaces a los documentos publicados en portales de la administración pública.

Dentro de los bancos de datos jurídicos de origen público, hallamos dentro del marco del sistema argentino de información jurídica (SAIJ), Infojus, que consiste en un servicio de información jurídica administrada por la Dirección Nacional del Sistema Argentino de Información Jurídica, dependiente de la Secretaría de Justicia del Ministerio de Justicia y Derechos Humanos. Infojus brinda información a magistrados, abogados, docentes, estudiantes y al público en general⁴¹⁵.

Infojus se caracteriza por su gratuidad, es decir que el acceso a sus contenidos es totalmente libre desde cualquier computador personal, sin claves de usuario ni otras condiciones que limiten su uso. Adicionalmente, despliega su trabajo en todo el país, por conducto de una red virtual y un sistema de terminales de consulta distribuidas

⁴¹⁴ ÍNDICE FUNDAMENTAL DO DIREITO. (2014, Septiembre 9). Consultada en julio 10, 2014.
<http://www.realjus.com.br/dji/>

⁴¹⁵ SISTEMA ARGENTINO DE INFORMACIÓN JURÍDICA. INFOJUS. (n.d.). Consultada en julio 2, 2014.
<http://www.infojus.gob.ar/>

en todo el territorio nacional. Lo anterior, reconoce modelos tecnológicos, que incluso permite que personas con discapacidad la puedan consultar⁴¹⁶

El sistema de información Infojus, contiene todos los documentos jurídicos, por medio de la extensa y actualizada biblioteca jurídica de normas, jurisprudencia y doctrina interrelacionadas entre sí, del ámbito nacional y provincial; y se accede a través de un buscador especializado, desarrollado para el derecho argentino⁴¹⁷.

Infojus cuenta con más de ochocientos mil documentos jurídicos tomados de la fuente oficial y es la única base de datos del país que ofrece las leyes nacionales sancionadas desde 1853, los decretos nacionales generales y vigentes desde 1975 y todos los decretos reglamentarios de leyes vigentes, Leyes provinciales provistas por sus propios organismos de origen, una selección de sentencias de jurisprudencia del ámbito federal, nacional y provincial, artículos de doctrina de importantes autores e investigadores del derecho, dictámenes de la Procuración del Tesoro de la Nación. Además, normas del Mercosur y resoluciones generales de la AFIP. Toda la información es ordenada y actualizada diariamente⁴¹⁸.

También encontramos en Argentina el banco de datos de la secretaria de jurisprudencia (SJ) de la Corte Suprema de Justicia de la Nación. En dicha base de datos encontramos un campo de novedades, en donde muestran los fallos recientes

⁴¹⁶ *Ibídem.*

⁴¹⁷ *Ibídem.*

⁴¹⁸ *Ibídem.*

más destacados, ordenados por fecha, con una descripción de los temas tratados, que permite acceder al texto completo de la sentencia objeto de búsqueda⁴¹⁹.

La Secretaría de Jurisprudencia, incorpora a su página web las sentencias publicadas en los tomos 1 a 307, de la colección de fallos de la Corte, que corresponde a los años desde 1863 hasta 1985. A partir de 1994, se encuentran los fallos completos, que se pueden consultar por los campos de actor, texto y expediente⁴²⁰.

Desde 1863 hasta 1966, en la base de datos solo se encuentran sumarios; desde 1966 hasta 1994 la base contiene los fallos en formato imagen y desde 1994 hasta septiembre de 2012 en soporte digital. Adicionalmente, se puede localizar boletines de jurisprudencia cuyo objetivo es brindar información sobre determinados núcleos temáticos en los que se ha pronunciado la Corte, con el propósito de ilustrar acerca de la doctrina elaborada⁴²¹.

En este país se encuentra la Red Federal de Bibliotecas Parlamentarias REFBIPAR, la que cuenta con una base de datos que contiene normas jurídicas emanadas de las Legislaturas Provinciales, que participan de la red⁴²².

Esta base de datos la administra la Biblioteca del Congreso de la Nación, quien inicio en 1987 la informatización de la legislación nacional y provincial, que hasta ese momento se procesaba mediante fichas catalográficas que se archivaban en

⁴¹⁹ ARGENTINA. CORTE SUPREMA DE JUSTICIA. (n.d.). Consultada en julio 5, 2014.
<http://www.csjn.gov.ar/jurisprudencia.html>

⁴²⁰ *Ibídem*.

⁴²¹ *Ibídem*.

⁴²² ARGENTINA. RED FEDERAL DE BIBLIOTECAS PARLAMENTARIAS. REFBIPAR. (2006). julio 5, 2014.
<http://refbipar.bcn.gov.ar/>

catálogos convencionales. Del trabajo conjunto de las Direcciones de Referencia Legislativa y de Informática de la BCN, surgió el sistema LEGIS, el cual permitió la incorporación de la información referencial contenida en las disposiciones legales, su clasificación temática por medio de un tesauro desarrollado específicamente para el sistema y un análisis de las modificaciones realizadas sobre la legislación ya existente, formando de esta manera un esqueleto de disposiciones que permite actualmente, conocer la situación de los incisos contenidos en los artículos de cada una de las normas nacionales y provinciales en cualquier momento de la historia de la disposición⁴²³.

La base de datos LEGIS fue aumentando la cantidad de disposiciones, mientras que el sistema adelantaba, adaptándose a los cambios tecnológicos. En el año 2006, con la masificación de Internet, en la Biblioteca del Congreso se crea la Red Federal de Bibliotecas Parlamentarias y se modifica el sistema centralizado LEGIS, el cual se transforma en un sistema de carga descentralizada, dando la posibilidad a cada una de las bibliotecas parlamentarias de Argentina, a que por medio de internet, incorporen en el sistema la información legal correspondiente a su jurisdicción, colaborando en la construcción de una base de datos general. A su vez, el sistema permite al usuario ir al sitio web que contiene las publicaciones oficiales, donde se encuentra el texto completo de la disposición⁴²⁴.

De otra parte, este sitio web, nos permite observar que tiene un sistema fácil, completo, práctico y de gran utilidad para los usuarios interesados en el conocimiento de las normas argentinas.

⁴²³ *Ibíd.*

⁴²⁴ *Ibíd.*

Por otro lado encontramos el Boletín Oficial de la República Argentina, que es un medio oficial de divulgación de las normas que emiten los entes nacionales y provinciales⁴²⁵.

El modelo de consulta de este boletín no es complicado, se puede consultar por tipo, número y año de la norma en cuestión y cuenta con un servicio pago de alertas, que permite obtener información sobre temas específicos de derecho⁴²⁶.

La Dirección de Informática Jurídica dependiente de la Fiscalía de Estado de la Provincia de Córdoba produce información desde la constitución provincial hasta leyes, decretos, decretos ley. Adicionalmente, presentan un resumen normativo mensual, entre otros⁴²⁷.

Dentro de los bancos de datos de origen público, también encontramos en Argentina el Sistema Informático CODEX, que constituye el soporte informático del Ordenamiento y Sistematización de normas de la Provincia de Córdoba, desarrollado por la Dirección de Procedimientos Administrativos y Sistemas Normativos y aprobado por la Secretaría de la Función Pública.

CODEX, inicio en julio de 2010 un banco de datos con dieciocho mil trecientas resoluciones ministeriales e incorpora a partir de octubre de 2011, información

⁴²⁵ ARGENTINA. BOLETÍN OFICIAL. (n.d.). Consultada en julio 6, 2014.

<http://www.boletinoficial.gov.ar/Inicio/Index.castle>

⁴²⁶ *Ibíd.*

⁴²⁷ CÓRDOBA (ARGENTINA). DIRECCIÓN DE INFORMÁTICA JURÍDICA. (n.d.). Consultada en julio 6, 2014. Legislación Provincial: <http://web2.cba.gov.ar/web/leyes.nsf/fri?OpenForm>

referencial de más de cuatro mil Leyes y Decretos, y más de treinta mil Resoluciones Ministeriales, todos ordenados y sistematizados, relacionados de modo integral⁴²⁸.

Este sistema de información es más completo y complejo que el que desarrolla La Dirección de Informática Jurídica dependiente de la Fiscalía de Estado de la Provincia de Córdoba. La base de datos se puede consultar por cualquiera de los campos disponibles y para optimizar, ampliar o restringir la búsqueda, se pueden combinar varios campos entre sí, con operadores lógicos que incorpora, trayendo como resultado un listado de normas con su extracto, que pueden ordenarse cronológicamente y permite ver un detalle de la misma y sus relaciones con otras normas, dando la opción de ir al diario oficial, para consultar el texto completo⁴²⁹.

El Ministerio de Justicia y Derechos humanos, por conducto de su Departamento de Biblioteca y Centro de Documentación cuenta con un compendio de bancos de datos de Argentina, que mediante link va a los sitios web de cada uno de ellas⁴³⁰.

El Ministerio de Economía y Finanzas Públicas, en su Centro de Documentación de Información cuenta con un sistema llamado Infoleg, que permite consultar la base de datos por tipo de norma, número y año. El área de información legislativa fue creada en 1930, con continuidad en el tiempo y la dotación de nuevas tecnologías, que la ha convertido en una herramienta de consulta importante en materia legislativa.⁴³¹

⁴²⁸ CÓRDOBA (ARGENTINA). Codex. Sistema de Ordenamiento y Sistematización de Normas Provinciales. (2015, Enero 31). Consultada en julio 6, 2014. <http://codex.cba.gov.ar/Paginas/Busqueda.aspx>

⁴²⁹ *Ibíd.*

⁴³⁰ ARGENTINA. MINISTERIO DE JUSTICIA Y DERECHOS HUMANOS. (2015, Febrero). Consultada en julio 6, 2014. Departamento de Biblioteca y Centro de Documentación: <http://www.biblioteca.jus.gov.ar/jurisprudencia-argentina.html>,

⁴³¹ ARGENTINA. MINISTERIO DE ECONOMÍA Y FINANZAS PÚBLICAS. (2005). Consultada en julio 6, 2014. Infoleg. Información Legislativa y Documental.: http://infoleg.mecon.gov.ar/?page_id=122,

Entre las funciones que cumple Infoleg, están las de efectuar el servicio de apoyo documental y de información legislativa dispuestos a satisfacer las necesidades del servicio jurídico, de los asesores y de las demás dependencias del Ministerio, así como también del público en general. Realizar la búsqueda y compilación de las leyes, decretos, decisiones administrativas, resoluciones y disposiciones, conforme a las publicaciones oficiales, manteniendo actualizada la base de datos, así como también, cuando se requiera, de las normas positivas de antecedentes jurisprudenciales y doctrinarios que correspondan. Efectuar los sistemas de búsqueda y compilación de las discusiones parlamentarias desarrolladas en el Congreso, relativos a los proyectos económico y finalmente coordina las actividades en materia de su competencia con organismos nacionales a efectos de obtener y brindar la información necesaria, para el cumplimiento de funciones comunes⁴³².

En conclusión, Infoleg presta servicios de consulta y asistencia documental, búsquedas especializadas en bases de datos legislativas y extranjeras, por conducto de internet; búsquedas por medio de la intranet del MECON; enlaces a bases de datos externas, capacitación y asistencia técnica, entre otros.

En Argentina también existen bancos de datos jurídicos de origen privado, entre los cuales podemos mencionar a *TodoelDerecho.com*, que es un portal jurídico, que utiliza como plataforma tecnológica internet. En dicho portal, se encuentran los textos completos de las normas económicas, impositivas y laborales. Adicionalmente, se halla legislación aduanera, comercial, previsional, societaria, bancaria y de distintos temas de interés de las provincias y del gobierno de la ciudad

⁴³² *Ibíd.*

de Buenos Aires. Este sitio web, contiene más de quinientos mil documentos jurídicos, incluyendo los últimos fallos y normas emanadas por los distintos organismos, dispuestos en cuatro bases de datos: legislación, con más de cincuenta mil normas con texto completo, del orden nacional y provincial ordenadas por año, tema y orden numérico; jurisprudencia, que contiene los fallos emitidos por los distintos tribunales del país, ordenados alfabéticamente, por tema, por tribunal y por materia; modelos de escritos, más de tres mil modelos de documentos con reglas generales, que se pueden adaptar a las necesidades, escritos judiciales, cédulas, cartas, etc., y un link de download, que permite descargar en el computador, el material más importante para el ejercicio profesional, leyes básicas, fallos, tratados internacionales, entre otros⁴³³.

Otra base de datos de origen privado, es la de informaticajuridica.com, cuyo objetivo es proveer información a investigadores, en el campo de derecho informático, en legislación, sentencias, master, seminarios, congresos y cursos, entre otros⁴³⁴.

Esta base de datos cuenta con más de veintidós mil doscientos artículos, de prensa, revistas, ponencias, mesas redondas, conferencias, jornadas, encuentros, seminarios, cursos, congresos y libros publicados sobre Derecho Informático e Informática Jurídica. Aparecen dos mil trescientos treinta dos autores de estos artículos y trabajos en temas de comercio electrónico, delitos informáticos, derechos de autor,

⁴³³ TODOELDERECHO.COM. (2014). Consultada en julio 6, 2014. <http://bases.todoelderecho.com>
Torres Andrade, C. A. (2008). Agentes y web semántica: elementos claves para el futuro del e-learning. *Revista Tecnura*, 11(22), 111-123.

⁴³⁴ INFORMÁTICA JURÍDICA.COM. (2015, Marzo 28). Consultada en julio 7, 2014. <http://www.informatica-juridica.com>

firma electrónica, nombres de dominio, protección de datos, trabajo y nuevas tecnologías⁴³⁵.

En igual sentido, los bancos de datos jurídicos *laleyonline.com*, del grupo Thompson Reuters, ofrece al usuario la posibilidad de consultar la revista jurídica Argentina, en donde se encuentra doctrina y jurisprudencia nacional, provincial y extranjera, con la viabilidad de tener texto completo y resumen. Asimismo se puede consultar los anales de legislación Argentina, que se refiere a la legislación nacional y provincial de todas las áreas del derecho y textos actualizados de los principales regímenes nacionales. Los fallos de la Corte Premium, su tratamiento diferencial y exclusivo de fallos de la Corte Suprema con el análisis de las vías de acceso al Tribunal, sentencias relacionadas de la Corte, tribunales extranjeros y organismos internacionales. Legislación Comentada Premium, como son los códigos civil, de comercio y penal, comentados y anotados, además de los principales regímenes nacionales, actualizados y vinculados temáticamente con legislación, doctrina y jurisprudencia. Antecedentes Parlamentarios, con el tratamiento parlamentario de las principales leyes nacionales. Doctrina judicial, con toda la información que el abogado litigante necesita para su labor diaria, así como cuenta con modelos contractuales y procesales, formularios y escritos judiciales, a más de una serie de recursos para la organización del estudio jurídico. Derecho del Trabajo, toda la información laboral y de la seguridad social: legislación, doctrina y jurisprudencia en un solo lugar con actualización en tiempo real, las principales leyes laborales comentadas, más de seiscientos convenios colectivos actualizados y correlacionados con sus escalas. Derecho internacional y del Mercosur, con la legislación vigente del

⁴³⁵ *Ibíd.*

Mercosur, derecho internacional penal, del trabajo e inmigración en Argentina, doctrina internacional y del Mercosur, además de jurisprudencia argentina e internacional y arbitraje internacional⁴³⁶.

En la categoría de los bancos de datos de origen privado encontramos otros como *informaticalegal.com.*, que consiste en una compilación de las principales normas relacionadas con la regulación de la informática, internet y las tecnologías de información en Argentina, desde el año 2000⁴³⁷.

También encontramos a *vlex.com*, información jurídica inteligente. En su comienzo ofreció información jurídica gratuita de España y América Latina. Posteriormente, a medida que se iban incorporando contenidos premium (libros, revistas, contratos), se desarrollaron productos de información jurídica por suscripción, en un primer momento en España y hoy en día se encuentra en varios países iberoamericanos. Dispone de una gran base de documentos jurídicos, doctrina con libros completos de autores reconocidos, fuentes normativas de jurisprudencia, desde el siglo XIX⁴³⁸.

En Argentina, también presta su servicio remunerado de información jurídica *microjuris.com*, es una plataforma digital de inteligencia jurídica, con soluciones integrales y a disposición más de cuarenta y cinco años de información jurídica. El portal funciona como un diario jurídico tradicional, en donde diariamente se puede consultar la normativa relevante a nivel nacional y provincial, en texto completo actualizada y vinculada; jurisprudencia a nivel nacional, federal y provincial,

⁴³⁶ *La Ley Online*. (2011). Consultada en julio 7, 2014. <http://www.laleyonline.com.ar>

⁴³⁷ *Informática Legal*. (2011). Consultada en julio 7, 2014. <http://www.informaticalegal.com.ar/legislacion-informatica/>

⁴³⁸ VLEX. ARGENTINA. (2014). Consultada en julio 7, 2014. <http://ar.vlex.com/jurisdictions/AR>

relacionada temáticamente con fallos en igual sentido y sentido contrario; doctrina de reconocidos autores; proyectos de ley, noticias jurídicas y cursos y actividades académicas. Finalmente, es importante destacar que este sitio se encuentra vinculado con las redes sociales⁴³⁹.

Por último, encontramos la base de datos jurídica argentinajudicial.com.ar, que se elabora y nutre con los aportes de los usuarios, quienes participan subiendo al sistema providencias y doctrina⁴⁴⁰.

En el ámbito académico se encuentra el sitio web Legislaw, es un sistema de información, que consiste en un directorio encaminado a facilitar la localización de sitios web con contenidos jurídicos.

Este sitio es administrado por profesionales y estudiantes y cuenta con vínculos de acceso preferente al resumen del día del Boletín Oficial de Argentina, a la guía judicial que publica el Ministerio de Justicia, a la base de datos legislativa Infoleg, obras y servicios públicos, a los proyectos legislativos en trámite en el Senado y en la Cámara de Diputados de la Nación, lo mismo que a los principales diarios del país. También se puede tener acceso desde la página principal a las novedades que publican la Corte Suprema de Justicia, el SAIJ, el Banco Central de Argentina, la AFIP y los concursos y quiebras del día anterior que publica el sitio web del diario judicial⁴⁴¹.

⁴³⁹ MICROJURIS.COM. INTELIGENCIA JURÍDICA. (1989-2015). Consultada en julio 7, 2014. <http://pr.microjuris.com/products-MJSuperior.jsp>

⁴⁴⁰ ARGENTINA JUDICIAL. (n.d.). Consultada en julio 10, 2014. <http://argentinajudicial.com.ar/>

⁴⁴¹ LEGIS LAW. (1999). Consultada en julio 10, 2014. www.legislaw.com.ar

La información jurídica, está contenida en diez canales y los canales correspondientes a Legislación, Jurisprudencia y Doctrina, están concebidos con la idea de permitir una evaluación del sitio, que es fuente de la información, y de facilitar el acceso inmediato a la misma, sin necesidad de recorrer distintas páginas hasta llegar a ella⁴⁴².

Los catálogos de legislaw, incluyen más de 600 sitios y cuenta con alrededor de 5.000 vínculos, que permiten ingresar a páginas en donde se localizan documentos, datos e información relevantes en el campo jurídico. Por conducto de ellos, legislaw facilita el acceso a sitios donde es posible hallar cientos de leyes nacionales y locales, fallos jurisprudenciales, opiniones doctrinarias, acceso a sistemas de autoconsulta de causas judiciales y administrativas, a bases de datos de interés práctico, a sitios gubernamentales y a muchas otras páginas de interés jurídico⁴⁴³.

2.15.16. Ecuador

En este país encontramos la base de datos contenida en la página principal del Consejo de la Judicatura, en el sitio web Función Judicial. El Departamento de Jurisprudencia e Investigaciones Jurídicas, cumple con unos de sus objetivos primordiales, al constituirse en un centro especializado de asesoramiento técnico – jurídico para la función judicial y una herramienta necesaria de consulta jurisprudencial para los servidores judiciales, profesionales del derecho, estudiantes y público en general, mediante la creación de bases de datos y de publicaciones periódicas que contribuyan a la conocimiento jurídico del país.

⁴⁴² Ibídem.

⁴⁴³ Ibídem

Para el cumplimiento de los anteriores objetivos, se creó el sistema de procesamiento de jurisprudencia, mediante la cual se accede a la jurisprudencia del más alto tribunal de justicia, Corte Nacional de Justicia, en donde por medio del módulo de búsquedas se puede consultar la información por el tesoro, número de juicio, partes procesales y número de resolución⁴⁴⁴.

El Departamento de Procesamiento de Jurisprudencia en un futuro brindará información sobre la línea jurisprudencial histórica y actual de la Corte de Casación; la reiteración de fallos, tres veces, que constituyen el sistema de precedente judicial obligatorio y vinculante; el criterio dirimente establecido por el pleno de la Corte Nacional de Justicia, sobre los fallos contradictorios; prontuarios; tesauros y repertorios de jurisprudencia⁴⁴⁵.

En Ecuador, también encontramos la página Derecho Ecuador, que permite consultar la legislación ecuatoriana: códigos, constitución de 2008, constituciones de la república, decretos presidenciales, derogatoria de disposiciones normativas, leyes ordinarias y leyes orgánicas. Así mismo permite buscar la jurisprudencia y ofrece la revista judicial que contiene diecisiete leyes orgánicas, siete códigos y setenta y seis leyes, entre estas incluida la nueva ley reformativa para la equidad tributaria en el Ecuador, el código tributario y ley de régimen tributario interno con sus correspondientes reformas, así como los proyectos de ley que están siendo

⁴⁴⁴ ECUADOR. CONSEJO DE LA JUDICATURA. (2011). Consultada en julio 14, 2014.
<http://www.funcionjudicial.gob.ec/>

⁴⁴⁵ *Ibíd.*

analizados y discutidos en la Asamblea Nacional Constituyente. Adicionalmente, en otro ícono presenta la posibilidad de consultar el derecho penal⁴⁴⁶.

La biblioteca legal en línea, de la empresa superley, en su sitio web, ofrece la posibilidad de consultar legislación, jurisprudencia, doctrina jurídica y tratados. La información se puede obtener por conducto de medios impresos, discos compactos, telefonía celular e internet. Adicionalmente, permite hacer enlaces con la Presidencia de la República, el Congreso Nacional y los fallos de la Corte Suprema de Justicia⁴⁴⁷.

Dentro de las empresas que prestan servicios de información legal en Ecuador, encontramos a ediciones legales, empresa que publica toda la legislación ecuatoriana, actualizada diariamente; jurisprudencia y la revista novedades jurídicas, entre otros. Su producto fielweb, es el único en el país, que cuenta con una versión propia y específica para iPad, iPhone, iPod, BlackBerry 4.5 y sistemas Android⁴⁴⁸.

También se encuentra el sistema integrado de la legislación ecuatoriana esilec, que almacena y organiza todos los actos normativos y administrativos, y toda la jurisprudencia de última instancia que se ha publicado en instrumentos oficiales, desde inicios de la República. Cuentan con más de doscientos sesenta y cinco mil normas, cada una de ellas registrada y catalogada desde su fecha de promulgación, con verificación diaria y determinación del periodo de vigencia, señalando los datos de las reformas que cada norma ha tenido⁴⁴⁹.

⁴⁴⁶ DERECHO ECUADOR. (2013). Consultada en julio 15, 2014. <http://www.derechoecuador.com>

⁴⁴⁷ SUPERLEY. INFORMACIÓN LEGAL AL DÍA. (n.d.). Consultada en julio 15, 2014. <http://www.superley.ec/super/>

⁴⁴⁸ EDICIONES LEGALES. (n.d.). Consultada en julio 15, 2014. <http://www.edicioneslegales.com.ec/>

⁴⁴⁹ LEXIS. (2015). Consultada en julio 15, 2014. <http://goo.gl/zzmG9B>

Esilec tiene características importantes, como contener todas las normas publicadas, organizadas por temas, materias, fecha, tipo, número y estado, entre otros. En cada norma se indica si se encuentra vigente, reformada o derogada. Sus textos se encuentran actualizados con las referencias que correspondan e igualmente, las normas incluyen concordancias y se hallan vinculadas con la jurisprudencia. Además, el sistema de esilec, ofrece la posibilidad de incorporar observaciones propias a los textos⁴⁵⁰.

2.15.17. México

De los bancos de datos jurídicos del pasado, en México encontramos el Instituto de Investigaciones jurídicas de la Universidad Nacional Autónoma de México. UNAM., con el sistema unan iure, cuyo objetivo es la compilación, análisis y sistematización de documentación legislativa y jurisprudencial, tanto nacional como extranjera, así como la realización de sistemas de información, consultables en disco compacto e Internet, capaces de proporcionar acceso a gran parte de la información nacional que se encuentra en la misma⁴⁵¹.

La información ha tenido un trabajo importante de recopilación, se ha realizado a lo largo de los sesenta años de existencia del Instituto de Investigaciones Jurídicas. Entre las diversas colecciones periódicas de carácter oficial se destacan el Diario Oficial de la Federación (*desde 1890*), todas y cada una de las colecciones de periódicos, gacetas y boletines oficiales de las treinta y dos entidades federativas (*desde 1930*) y el Semanario Judicial de la Federación (*desde 1870*).

⁴⁵⁰ LEXIS. INTELIGENCIA JURÍDICA. (1998). Consultada en julio 15, 2014. <http://www.lexis.com.ec>

⁴⁵¹ INSTITUTO DE INVESTIGACIONES JURÍDICAS. (1995). Consultada en julio 20, 2014. <http://www.juridicas.unam.mx/>

Adicionalmente se puede consultar la legislación mexicana desde el siglo XVIII, diversos compendios de legislación local publicados por gobiernos estatales, legislación extranjera, jurisprudencia del Tribunal Fiscal de la Federación, criterios de interpretación del Tribunal Superior de Justicia del Distrito Federal, índices de jurisprudencia elaborados por el propio Centro, así como jurisprudencia de países como Colombia, Costa Rica, Cuba, Chile, Ecuador, El Salvador, España, Estados Unidos, Francia, Guatemala, Honduras, Israel, Italia, Nicaragua, Panamá, Perú y Venezuela. El Centro cuenta con una sala de consulta en donde los investigadores del Instituto y público en general pueden consultar la documentación legislativa y jurisprudencial⁴⁵².

El sistema unam-jure fue el primer sistema de información desarrollado por el instituto de investigaciones jurídicas, que lo inició desde hace dos décadas y lo diseñó para para almacenar, identificar y dar seguimiento a las disposiciones legislativas emitidas en la República Mexicana, publicadas en el Diario Oficial de la Federación y en todos los periódicos, gacetas y boletines oficiales de los Estados.

Este sistema, contiene fichas de análisis que ofrecen información sobre la expedición y reforma de leyes, códigos, decretos, tratados, constituciones, estatutos, normas oficiales, bandos, ordenanzas, planes, presupuestos, reglamentos, tarifas, manuales, contratos ley, circulares, avisos y acuerdos, entre otros, publicados en todo el

⁴⁵² *Ibíd.*

territorio nacional. En su momento el unam-jure llegó a constituirse como el banco de información legislativa más importante del país⁴⁵³.

En la década de los noventa el instituto de investigaciones jurídicas, por conducto de su centro de documentación y el departamento de sistemas ha desarrollado diversos sistemas que proporcionan información jurisprudencial en derecho penal, civil, laboral, constitucional, así como información legislativa federal y estatal, en texto completo. Estos sistemas se pueden consultar por conducto de disco compacto e internet y en ocasiones para su consulta interna en entidades gubernamentales⁴⁵⁴.

El Instituto, también ha desarrollado otros sistemas de información entre los que podemos mencionar el CFE-JURE, DF-JURE, sistema de información jurídica en materia de seguridad pública, UNIJUS, ADPROJUS, sistema de información jurídico laboral, sistema de información jurídico Constitucional, JUS-TUR, sistema de información jurídico electoral e ISSSTE-JURE.

En México se aprecia la base de datos de legislación federal mexicana (BDLFM), creada por el Instituto Mexicano de Estrategias, que a diario recoge la información del diario oficial de la Federación, que inciden sobre las políticas de gobierno. Hasta la fecha se han incluido más de quince mil ordenamientos jurídicos, y cada uno de ellos se ha catalogado, para poderlo localizar e interconectar con el resto de la legislación⁴⁵⁵.

⁴⁵³ *Ibíd.*

⁴⁵⁴ *Ibíd.*

⁴⁵⁵ INSTITUTO MEXICANO DE ESTRATEGIAS S.C. (n.d.). Consultada en junio 23, 2015. <http://www.siemexico.mx/>

BDLFM, se caracteriza por el uso de tecnología de punta y la metodología ideada para correlacionar leyes por conceptos, relaciones funcionales, organización jerárquica, evolución en el tiempo y vinculación territorial de la legislación federal mexicana. Utilizan tesauros de palabras clave, lo que significa que cuentan con un diccionario de conceptos interrelacionados entre sí⁴⁵⁶.

La base de datos de legislación federal mexicana permite buscar la información para conocer la evolución y vigencia del ordenamiento. Asimismo el banco de datos identifica los antecedentes, objetivos, constantes cambios de la legislación y su vinculación con otros instrumentos de gobierno⁴⁵⁷.

La búsqueda de la legislación se puede hacer mediante el módulo de interconexiones, que permite identificar la información que necesita, el lugar que ocupa en las políticas de gobierno y sus vínculos con el sistema jurídico mexicano. Igualmente se puede hacer la consulta mediante el módulo de palabras claves, a partir de conceptos específicos. Se encuentra en etapa de construcción el módulo de análisis, cuyo fin es el de proporcionar información adicional de gran importancia, sobre legislación seleccionada previamente⁴⁵⁸.

Otro banco de datos importante es el de la Cámara de Diputados de México. El usuario puede tener acceso a los textos completos, correspondientes a las leyes federales vigentes, reformas a la constitución, sumarios de reformas, reformas por legislatura, marco jurídico del congreso, leyes federales abrogadas, reglamentos de

⁴⁵⁶ *Ibíd.*

⁴⁵⁷ *Ibíd.*

⁴⁵⁸ *Ibíd.*

leyes federales, normas reglamentarias, leyes y poderes estatales e índices del diario oficial, entre otra normatividad actualizada. De otro lado cuenta con varias compilaciones temáticas y permite realizar búsquedas, cuyos resultados son generales⁴⁵⁹.

Otras bases de datos que se encuentra en México son los bancos de datos especializados en línea, de la Suprema Corte de la Justicia de la Nación, que coadyuvan para satisfacer las necesidades de información y complementar el acervo bibliohemerográfico bajo resguardo del Sistema Bibliotecario de la Suprema Corte de Justicia de la Nación, al permitir encontrar y recuperar documentos novedosos tanto de contenido jurídico como de materias relacionadas. El acceso a ellas se hace mediante cualquiera de las bibliotecas del alto tribunal.

El objeto de estas bases de datos es ampliar los recursos disponibles en las bibliotecas a todos los miembros del poder judicial y de la sociedad en general, que podrán acceder a las mismas, mediante un nombre de usuario y/o contraseña, suministrada por cualquier integrante de la red de bibliotecas del poder judicial de la federación⁴⁶⁰.

En México encontramos las bases de datos de JURIVERSIA, que ofrece información jurídica para Brasil, Colombia, México, Perú y Venezuela. En el caso de México la base de datos está conformada por legislación federal, estatal, municipal y el diario oficial de la federación, así como con jurisprudencia de la Suprema Corte de Justicia,

⁴⁵⁹ MÉXICO. CÁMARA DE DIPUTADOS. (2006). Consultada en julio 21, 2014.
<http://www.diputados.gob.mx/LeyesBiblio/index.htm>

⁴⁶⁰ MÉXICO. CORTE SUPREMA DE JUSTICIA. (n.d.). Consultada en julio 21, 2014.
<https://www.scjn.gob.mx/Paginas/Inicio.aspx>

Consejo de la Judicatura Federal, Tribunal Electoral, Tribunal Federal de Justicia Fiscal y tribunales colegiados. Tanto la legislación como la jurisprudencia son desplegadas con un índice cronológico⁴⁶¹.

2.15.18. Perú

En dicho país encontramos el Sistema Peruano de Informática Jurídica, SPIJ, es una base de datos del Ministerio de Justicia y Derechos Humanos, contiene la legislación, con textos completos, concordados y actualizados. Así mismo, el banco de datos contiene jurisprudencia judicial y administrativa, directorio de asesorías jurídicas y aquella información jurídica complementaria⁴⁶².

La búsqueda y recuperación de la información a nivel de documento jurídico legislación se puede obtener mediante palabras, frases, número y tipo de norma, materia, sectores y fecha de publicación, entre otros. Igualmente permite saber el estado de la norma, si ha sido modificada, derogada y las concordancias de la misma. De igual manera, admite la impresión de todos los textos, formatos, anexos y gráficos que se encuentren en el sistema y la copia a un procesador de palabra⁴⁶³.

Asimismo en el Perú se encuentra el archivo digital de la legislación de este país, desarrollado por el Congreso de la República, corresponde a un sistema indexado de los documentos digitalizados de la legislación, decretos y normas, mediante técnicas de reconocimiento óptico de caracteres (OCR), admite recuperar y consultar la

⁴⁶¹ JURIVERSIA. (n.d.). Consultada en julio 21, 2014. <http://www.juriversia.com/Mexico/index.aspx>

⁴⁶² SISTEMA PERUANO DE INFORMACIÓN JURÍDICA. SPIJ. (n.d.). Consultada en julio 22, 2014. <http://spij.minjus.gob.pe>

⁴⁶³ *Ibíd.*

totalidad del documento en línea. La búsqueda de las leyes se puede obtener mediante número, dación, promulgación, texto y permite definir en cada uno de ellos las palabras claves de cada norma, para recuperar el documento digital de la ley⁴⁶⁴.

En el diario el peruano-boletín oficial, se puede consultar las normas legales, entre las que se encuentran normas publicadas en el día, separata de estas normas, sentencias en casación, jurisprudencia, procesos constitucionales, declaraciones juradas y el boletín oficial. Las búsquedas se pueden realizar por la fecha de publicación del documento jurídico y permiten la recuperación del documento en digital⁴⁶⁵.

2.15.19. Costa Rica

En Costa Rica se encuentra el Sistema Costarricense de Información Jurídica, SCIJ, nace del Programa de Modernización de Administración de Justicia, desarrollado mediante contrato entre Banco de Desarrollo Interamericano BID y el gobierno de la República de Costa Rica⁴⁶⁶.

El SCIJ, ofrece consulta de información normativa y jurisprudencial en forma integrada. Contiene la legislación promulgada desde el año de 1821, entre lo que podemos mencionar a las leyes, decretos ejecutivos, convenios internacionales,

⁴⁶⁴ PERÚ. CONGRESO DE LA REPÚBLICA. (n.d.). Consultada en julio 26, 2014. <http://www.congreso.gob.pe/>

⁴⁶⁵ EL PERUANO. DIARIO OFICIAL. (2010). Retrieved Julio 26, 2014, from www.elperuano.com.pe, Especificación HTML 4.01. (n.d.). Consultado en marzo 7, 2015. ConClase.net

⁴⁶⁶ SISTEMA COSTARRICENSE DE INFORMACIÓN JURÍDICA. SCIJ. (n.d.). Consultada en junio 23, 2015. Sistema Costarricense de Información Jurídica, SCIJ: <http://www.pgrweb.go.cr/scij/>

reglamentos y demás normas de aplicación general. Igualmente se puede consultar la jurisprudencia de Tribunales, Salas de Casación y Sala Constitucional de Corte Suprema de Justicia; la jurisprudencia administrativa de la Procuraduría General de la República y del Ministerio de Hacienda⁴⁶⁷.

La búsqueda de la información se puede hacer mediante tesauros, índice alfabético de temas, por temas y subtemas, por datos de la sentencia, por despacho, por año y búsquedas libres⁴⁶⁸.

También encontramos en este país la base de datos del centro de información jurídica en línea, conocido como CIJUL en línea, forman parte integral de él, la Facultad de Derecho de la Universidad de Costa Rica, la Biblioteca de Derecho de la Universidad de Costa Rica y la Biblioteca del Colegio de Abogados. CIJUL en Línea ofrece sus servicios en forma personalizada a todos aquellos profesionales en derecho que se encuentren al día en su colegiatura y que soliciten una cuenta de usuario, con la cual podrán tener acceso en línea a las investigaciones realizadas por el centro, solicitar nuevas investigaciones y consultar las fuentes de información disponibles en sus bases de datos referenciales⁴⁶⁹.

Entre los objetivos que cumple CIJUL en línea está la de ofrecer a los profesionales en derecho acceso en línea a diferentes fuentes bibliográficas de información jurídica, contribuyendo así al desarrollo profesional de los mismos; mantener un fondo bibliográfico actualizado tanto en formato impreso como digital en materia

⁴⁶⁷ *Ibídem*.

⁴⁶⁸ *Ibídem*.

⁴⁶⁹ *CIJUL en Línea*. CENTRO DE INFORMACIÓN JURÍDICA EN LÍNEA DE COSTA RICA. (2014). Consultada en julio 29, 2014. <http://cijulenlinea.ucr.ac.cr/>

jurídica, para ofrecer servicios eficientes que faciliten el ejercicio de la profesión; brindar servicios actualizados de información, en forma personalizada, de acuerdo con las necesidades y requerimientos de los profesionales en esta área; brindar asesoramiento y capacitación a los abogados en el uso de herramientas tecnológicas; búsqueda efectiva de fuentes bibliográficas y promover la capacitación virtual en materia jurídica⁴⁷⁰.

El Centro hace uso de sistemas y tecnologías de punta en información y comunicación (TICs) y ofrece sus servicios por conducto de internet, vía telefónica o en forma presencial en sus instalaciones. Adicionalmente en el centro se puede reproducir libremente las constituciones, leyes, decretos, acuerdos municipales, reglamentos y demás actos públicos⁴⁷¹.

2.15.20. Puerto Rico

El portal oficial del Estado Libre Asociado de Puerto Rico⁴⁷², ofrece un sin número de servicios de búsqueda entre los que se encuentran: la búsqueda por casos que se verán en las salas de los tribunales de Puerto Rico (remitiendo al sistema de búsqueda de la Rama Judicial de este país que será explicada más adelante), la búsqueda Registro Contratistas (DACO), la búsqueda Registro de Marcas de Fábrica, etc.

⁴⁷⁰ Ibídem.

⁴⁷¹ Ibídem.

⁴⁷² ESTADO LIBRE ASOCIADO DE PUERTO RICO. (2013).

<http://www2.pr.gov/Pages/servicefilter.aspx?cat=servicios&type=ciudadano>

En el caso de la Rama Legislativa de ésta Nación, las búsquedas web de leyes, resoluciones de aprobación y medidas de fortaleza, cuentan con una base de datos especializada, cuyo objetivo, es satisfacer las necesidades de información legal, bajo la garantía del sistema bibliotecario del Congreso. En su página, el cuerpo legislativo facilita las búsquedas a través de un prontuario de categorías⁴⁷³, en el que se pueden realizar las búsquedas según la vigencia de las leyes (leyes nuevas, enmendadas y derogadas) y el año de promulgación, utilizando para estos efectos, el instrumento lingüístico de tesaurus jurídico, que representa de manera uniforme el contenido de los documentos que integran las colecciones documentales del órgano legislador. Para concluir, cabe resaltar que las leyes pueden ser consultadas tanto en español como en inglés.

La rama judicial de Puerto Rico dispone de una base de datos, en la que se puede recuperar los pronunciamientos de las Cortes, mediante un sistema de búsqueda por caso específico, fecha o tema⁴⁷⁴.

La base de datos de la Rama Judicial, a pesar de enseñar instrucciones claras, su interface es antigua y poco agradable a la vista. La búsqueda se hace compleja, ya que arroja excesivos resultados, siendo la mayoría de poca relevancia y utilidad.

Asimismo, en Puerto Rico se encuentra el sitio web microjuris.com, ofrece colecciones que contienen información legal digital, entre otros, leyes, jurisprudencia local y federal, reglamentos, tratadistas y doctrina. Encarna una herramienta de

⁴⁷³ OFICINA DE SERVICIOS LEGISLATIVOS DE LA ASAMBLEA LEGISLATIVA DE PUERTO RICO. (n.d.) Consultado en junio 23, 2015. Prontuario: <http://www.oslpr.org/prontuario/>

⁴⁷⁴ LA RAMA JUDICIAL DE PUERTO RICO. (n.d.). Consultada en agosto 5, 2014. <http://www.ramajudicial.pr/>

investigación o trabajo digital en línea y en tiempo real, que se ajusta muy bien al sistema legal de Puerto Rico⁴⁷⁵.

Este sitio dispone de un motor de búsqueda, que da la opción de filtrar la información por colección; desde tribunales, artículos, revistas y distintas cortes; compendios, tales como laboral, ambiental, contributiva; periodos y la posibilidad de ordenar los resultados que arroje el buscador.

De igual manera, en Puerto Rico se encuentran portales de internet de acceso gratuito, es así como el colegio de abogados, facilita la consulta del texto completo de las avanzadas del Tribunal Supremo, desde 1998 hasta la fecha y para encontrar la información es necesario saber las partes, el año y el mes de la decisión. Adicionalmente contiene varios reglamentos e índices de temas⁴⁷⁶.

De acceso gratuito también se halla la base de datos del departamento de Estado, que contiene las órdenes ejecutivas, desde 1996. Por medio de esta página se puede conectar con el servicio de lexisnexis y consultar las leyes de Puerto Rico.

Igualmente el portal de LexisNexis permite la consulta gratuita de las leyes de Puerto Rico en español e inglés. LexJuris, también de acceso gratuito, facilita la consulta de la legislación y jurisprudencia de este país desde 1997, e incluye una selección de leyes de uso frecuente, tales como, ley electoral, código civil, código de comercio y otros. Asimismo, encontramos a la oficina de servicios legislativos, que ofrece una

⁴⁷⁵ MICROJURIS.COM. INTELIGENCIA JURÍDICA. (1989-2015). Consultada en agosto 5, 2014.
<http://pr.microjuris.com/products-MJSuperior.jsp>

⁴⁷⁶ DERECHOUPR.COM: tu Shortcut a la Investigación Jurídica. (2015). Consultada en agosto 5, 2014.
<http://derechoupr.com>

de las bases de datos más recientes e importantes de Puerto Rico, de acceso gratuito, disponible para el público en general y contiene la base de datos del sistema de interrogación del trámite legislativo. PRLEX, contiene las decisiones del Tribunal Supremo desde 1998 hasta la fecha y permite consultar en forma sencilla. Las decisiones están organizadas por número, partes y temas. De acceso gratuito, el Tribunal Supremo de Puerto Rico, brinda sus decisiones desde 1998 hasta la fecha. Incluye el número de la decisión, del caso, la fecha y el tipo de opinión⁴⁷⁷.

2.15.21. Uruguay

En Uruguay encontramos a inforlex, base de datos jurídica que contiene jurisprudencia y doctrina en todas las materias, resúmenes de más de cuarenta años. En cuanto a legislación, posee acápites desde 1830, texto completo desde 1935, decretos desde 1966, códigos y resoluciones⁴⁷⁸. Es una base de datos que requiere de suscripción.

Otra base de datos que se encuentra en Uruguay es la Ley de Thomson Reuters. Para el ingreso a dicho banco de datos se requiere suscripción, pero la compañía dispone de un demo interactivo, que permite observar la estructura de la base de datos⁴⁷⁹.

La ley cuenta con la revista la ley Uruguay online, contiene jurisprudencia, doctrina nacional y extranjera, en texto completo y resúmenes. Adicionalmente, la justicia Uruguay en línea contiene jurisprudencia y doctrina nacional, en texto completo y

⁴⁷⁷ *Ibíd.*

⁴⁷⁸ INFORLEX. SISTEMA DE ÍNDICES JURÍDICOS. (1993-2013). Consultada en agosto 26, 2014. <http://www.inforlex.com.uy/>

⁴⁷⁹ LA LEY ONLINE. (2011). Consultada en agosto 26, 2014. <http://www.laleyonline.com.uy>

resúmenes. La búsqueda de la información se puede hacer mediante los temas, para lo cual desarrollan tesauros jurídicos, por palabra libre y mediante una búsqueda avanzada⁴⁸⁰.

Entre los bancos de datos que se localizan en el Uruguay, que requieren de una suscripción son las bases de datos Cade Uruguay, que reúne en un solo producto y un solo proveedor la información sobre todas las ramas del derecho. Ofrece la legislación, jurisprudencia y doctrina, más de cien mil documentos, en texto completo, con enlaces entre ellos y con índices bibliográficos, calculadoras, agenda de procuración, programas utilitarios, videoconferencias, cotizaciones de monedas, índices y coeficientes de interés, formularios notariales y de presentación de la demanda, escritos forenses⁴⁸¹.

En el módulo jurídico de las bases de datos Cade se encuentra la normativa, tales como las constituciones, códigos, textos ordenados, leyes desde 1900 hasta la fecha, decretos, resoluciones del poder ejecutivo y circulares de la Suprema Corte de Justicia. Adicionalmente ofrece jurisprudencia, tales como las sentencias de la Suprema Corte de Justicia, del Tribunal de lo Contencioso Administrativo, de los Tribunales de Apelaciones Civil, de Familia, del Trabajo y Penal, de los Juzgados Letrados y sentencias de los Juzgados de Paz⁴⁸².

La Contaduría General de la Nación facilita el acceso gratuito a la base de datos jurídica de Administración Pública Presupuestal, resultado del trabajo conjunto con

⁴⁸⁰ *Ibíd.*

⁴⁸¹ CADE. *Cuatro Décadas Brindando Servicios de Alta Calidad*. (2012). Consultada en agosto 26, 2014. www.cade.com.uy/

⁴⁸² *Ibíd.*

IMPO, centro de información oficial. La base de datos mencionada, tiene como objetivo poner al alcance del usuario información preseleccionada de la normativa referente a la materia, en forma rigurosa, la que es actualizada en forma permanente⁴⁸³.

2.15.22. Panamá

Los bancos de datos jurídicos en internet que se encuentran en Panamá, uno de los más importantes es el que ofrece la Procuraduría de la Administración Infojurídica IJ, que corresponde a un sistema de información de normas y afectaciones panameñas, las que se encuentran con anotación de vigencias e inconstitucionalidades, en texto completo, desde 1903 hasta la fecha. Inforjurídica brinda por conducto de su base de datos la posibilidad de consultar leyes, decretos de gabinete, resoluciones, reglamentos, decretos leyes, decretos ejecutivos, acuerdos y demás normas de carácter general. Adicionalmente, se puede tener acceso a todas las Gacetas Oficiales publicadas por la Asamblea Nacional, en formato PDF⁴⁸⁴.

Infojurídica nace de la necesidad que tiene la Procuraduría de la Administración de cumplir con la atribución legal de certificar la vigencia de las normas legales del país. Es importante anotar que la búsqueda de la información se puede realizar por temas, lo que nos indica que desarrollan instrumento lingüístico, para la normalización del lenguaje⁴⁸⁵.

⁴⁸³ URUGUAY. CONTADURÍA GENERAL DE LA NACIÓN. (n.d.). Consultada en agosto 26, 2014.
<http://goo.gl/vnKr8L>

⁴⁸⁴ PANAMÁ. PROCURADURÍA DE LA ADMINISTRACIÓN. INFOJURÍDICA. (2012). Consultada en agosto 27, 2014.
<http://goo.gl/JV7UTI>

⁴⁸⁵ Ibídem.

En Panamá también encontramos el banco de datos que ofrece el Tribunal electoral, que cuenta con información relacionada con publicaciones, boletines electorales y jurisprudencia, relacionada con el tema de su competencia⁴⁸⁶.

2.15.23. África Del Oeste

El primer banco de datos jurídico en África del Oeste se encuentra en el Centro Nacional de Documentación Jurídica, cuyo sitio web es <http://www.cndj.ci/>. El Centro mencionado, es una institución pública, adjunto al Ministerio de Justicia, se constituyó en 1995 y oficialmente abrió sus puertas al público en 1997, cuyo objetivo es la creación de un banco de datos, para la consulta, publicación y difusión de cualquier texto legal⁴⁸⁷.

El banco de datos legislativo del Centro Nacional de Documentación Jurídica, contiene más de 14.000 artículos, y es actualizada regularmente. La de jurisprudencia contiene más de 7000 decisiones emitidas tanto por los Tribunales de Costa de Marfil como por el Tribunal Común de Justicia y Arbitraje, CCJA⁴⁸⁸.

Después de hacer un recorrido de los bancos de datos jurídicos que existen en los diferentes países del mundo, es importante mencionar que encontramos sitios web, que son proveedores globales de información jurídica y ofrecen bancos de datos en

⁴⁸⁶ PANAMÁ. COMISIÓN NACIONAL DE REFORMAS ELECTORALES. (n.d.). Consultada en agosto 27, 2014. <http://www.tribunal-electoral.gob.pa/html/>

⁴⁸⁷ CENTRE NATIONAL DE DOCUMENTATION JURIDIQUE. (2012-2015). Consultada en septiembre 1, 2014. <http://www.cndj.ci/index.php>

⁴⁸⁸ *Ibíd.*

distintos países, en idiomas diferentes. Es así como encontramos a badaj.org, vlex.com, iate.europa.eu, icrc.org, Cicr, entre otros.

De acuerdo con lo anterior, podemos concluir que en casi todos los países antes mencionados, la información jurídica, en sus tres niveles, legislación, jurisprudencia y doctrina, está a disposición, para consulta en bancos de datos, independientemente que la información la presenten de diferente manera, dependiendo del país; así, la consulta de la información, en unos se necesite suscripción y en otros se pueda realizar en forma gratuita y libre.

3. BANCO DE DATOS JURÍDICOS: FUTURO

3.1. La era de la sobreinformación y el Big Data

Como bien se pudo apreciar en la segunda parte, existe en el mundo una gran cantidad de bases de datos jurídicas, cada uno de ellas contiene a su vez miles de documentos, de ahí que el usuario se vea obligado a analizar dentro de la lista de resultados de búsqueda, cuáles son los más pertinentes.

Las bases de datos contienen innumerables cantidades de documentos jurídicos, esto trae como consecuencia una dificultad en la selección oportuna y pertinente de los mismos. A pesar de que las bases de datos tengan por objetivo facilitar la búsqueda y organizar la información, el usuario se encuentra hoy en día con una nueva tarea: seleccionar dentro de los resultados que arroja la búsqueda los documentos pertinentes y analizarlos. Para cualquier operador jurídico estas tareas se tornan dispendiosas e impiden utilizar un mayor tiempo para el análisis del caso concreto.

Se evidencia un nuevo reto de las bases de datos jurídicas, que consiste en ofrecer al usuario un análisis previo de la información con resultados de búsqueda aún más pertinentes y actualizados en la mayor medida de lo posible, para así presentar de forma más efectiva la información dentro del cúmulo de documentos jurídicos que se insertan dentro de ella.

Existe tanta cantidad de información dentro de una base de datos, que este fenómeno puede ocasionar el efecto inverso en el objetivo de las mismas, que

consiste en organizar y proporcionar información. Nos encontramos así frente a la era de la sobreinformación.

Este fenómeno no se presenta únicamente dentro de los bancos de datos jurídicos, sino dentro de todas las bases de datos, cualquiera sea su naturaleza. En consecuencia, en la actualidad el surgimiento del *big data* empieza a revolucionar la forma en que se realiza el tratamiento de la información.

La manifestación del Big Data o de los grandes volúmenes de datos, se constituye en tema de gran interés, en la actualidad, lo que ha sido definido por varios autores, destacándose la definición de McKinsey Global Institute, que sostiene que el “big data se refiere a los conjuntos de datos cuyo tamaño está más allá de las capacidades de las herramientas típicas de software de bases de datos para capturar, almacenar, gestionar y analizar”⁴⁸⁹.

De acuerdo con el G20, en su dictamen de marzo de 2013, señala que “el *Big Data* hace referencia a un crecimiento exponencial de la disponibilidad y de la automaticidad en el uso de la información: esto hace referencia a las gigantes bases de datos que poseen las corporaciones, el gobierno y todas las demás organizaciones, que luego son analizadas ampliamente utilizando algoritmos informáticos”⁴⁹⁰.

⁴⁸⁹ MANYIKA, J., CHUI, M., BROWN, B. & OTROS. (2011, Mayo). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute: <http://goo.gl/3rwzGX>

⁴⁹⁰ “Big data refers to the exponential growth both in the availability and the automated use of information: it refers to gigantic digital datasets held by corporations, governments and other large organisations, which are then extensively analysed (hence the name: analytics100) using computer algorithms”. *Article 29 Data Protection Working Party*. Opinion 03/2013. (2013, Abril 2). http://idpc.gov.mt/dbfile.aspx/Opinion3_2013.pdf

También, el Big Data está definido como la regla de las tres V: el volumen, la variedad y la velocidad⁴⁹¹. Lo anterior se refiere a que el contiene un volumen importante de datos que sobrepasa el límite posible para manejar los datos, en una solución de bases de datos relacional.

Es importante precisar que, cuando se habla del volumen del *Big Data*, mal se haría en afirmar que este consiste en sobrepasar la capacidad de almacenamiento o de software existente en la actualidad. Primero, porque esta definición podría limitarse en el tiempo, teniendo en cuenta que la tecnología avanza a pasos agigantados y que esta evoluciona constantemente en lo que se refiere a almacenamiento. En segundo lugar, porque el concepto de *big data* no reside únicamente en el hecho de la existencia de grandes volúmenes de información sino en el fenómeno que se ocasiona con su surgimiento, es decir, mal se haría en definir el *Big Data* como un conjunto de datos que excede la capacidad de almacenamiento.

Es posible percibir, a diario, que el crecimiento exponencial de los datos brutos, que surgen gracias a las diferentes fuentes de información y que incluso pueden llegar a sobrepasar las capacidades de tratamiento de la misma, dificulta realizar un análisis de aquella, en tiempo record. De ahí que la velocidad sea la segunda característica del *Big Data* y que la rapidez en el tratamiento de la información, en una base de datos voluminosa, pueda ser efectuada incluso en tiempo real, para así poder capturar la masa de datos en movimiento.

⁴⁹¹ COMMISSION NATIONALE DE L'INFORMATIQUE ET DES LIBERTÉS. (n.d.). Vie privée à l'horizon 2020 Paroles d'experts. Cahiers Ip. Innovation & prospective, p. 18.

Ahora bien, la variedad de estos datos, es también una de las características del *Big Data*, toda vez que los formatos pueden ser múltiples, lo que permite encontrar el valor en el conjunto de datos accesibles, sean estos estructurados o no. Finalmente, hay quienes afirman que se debe agregar otra “V” a las características del *Big Data*: el valor de los datos, en la medida en que el *Big Data* transforma el dato en información⁴⁹².

Para concluir, el *Big Data* debe ser entendido como un fenómeno actual que consiste en la optimización de recursos para el manejo, organización y análisis de grandes cantidades de información, comprendidas dentro de una base de datos con contenido variable. La importancia del *Big Data* reside no solo en el volumen de los datos sino por la nueva forma en que son analizados.

El *big data* cuenta con un conjunto voluminoso y variado de datos, excediendo o no la capacidad de almacenamiento. La organización de estos datos tiene por objetivo seleccionar, analizar y procesar la información contenida en la base de datos, para arrojar así resultados pertinentes dentro de la búsqueda de forma inteligente, actualizada y en tiempo record. El gran reto de la existencia del *big data*, más allá del almacenamiento, consiste en la capacidad de explotación y de optimización del contenido de los datos. En otras palabras, el *Big Data* permite dejar hablar, los datos⁴⁹³.

⁴⁹² *Synthèse du marché Big Data*. (2013). Big Data, p. 3. París.

http://corpevents.fr/bigdata14/docs/Dossier_de_synthese_BIG_DATA_PARIS_2013.pdf

⁴⁹³ CUKIER, K., & MAYER-SCHOENBERGER, V. (2014). *Big Data: La révolution des données est en marche*. Robert Laffont. p. 14.

Un ejemplo que aplica el *big data*, es la *GDELT Global Data Base of Events Languages and Tones*, desarrollada por la Universidad de Georgetown y Kalev Leetaru, de Yahoo, que compila todas las actualidades (comunicados de prensa, artículos, discursos etc.), aparecidos desde 1979. Esta base de Datos, aplica técnicas de análisis semántico y algoritmos autodidactas para facilitar la comprensión de eventos recientes y de los principios causa y efecto para lograr predecir los conflictos mundiales. Se encuentra en *open source* con más de 250 millones de eventos⁴⁹⁴.

Ahora bien, este tema trasciende a las bases de datos jurídicos, toda vez que ellas no son ajenas al fenómeno del *Big Data*. Los bancos de datos jurídicos contienen grandes volúmenes de información. Para ilustrar dicha afirmación es suficiente examinar cuántos documentos jurídicos fueron producidos en un solo año y luego insertados en la base de datos jurídicos, de las altas cortes de Colombia. El número de providencias emitidas por estos cuerpos colegiados, durante el año 2014, corresponde a treinta y cinco mil sesenta y ocho documentos, relativos a providencias judiciales⁴⁹⁵.

La cantidad de información jurídica que se profiere, a diario, genera un nuevo reto: actualizarla información en tiempo record y, de ser posible, en tiempo real, con el fin de mejorar la toma de decisiones de los diferentes operadores jurídicos que utilizan las bases de datos. La velocidad del flujo, de esos datos, es una característica determinante para afirmar que no encontramos en frente del fenómeno del *Big Data*.

⁴⁹⁴ GDELT. GLOBAL DATA BASE OF EVENTS, LANGUAGES AND TONES. (2013-2014). Consultada en diciembre 3, 2014. <http://gdeltproject.org/>

⁴⁹⁵ RAMA JUDICIAL DE COLOMBIA. (2015). Estadísticas Judiciales Enero 1 a Diciembre 31 de 2014: <http://www.ramajudicial.gov.co/web/estadisticas-judiciales/movimiento-de-procesos-del-1-de-enero-a-31-de-diciembre-2014>

De otra parte, las fuentes de los documentos jurídicos son variadas, a pesar que en su gran mayoría son textos; también se incluyen videos de audiencias, coloquios, comunicados de prensa, con diferentes versiones de software. Incluso, un banco de datos jurídico podría incluir flujos de clics, para examinar los intereses particulares de cada usuario.

El fenómeno del volumen, la velocidad y la variedad de las bancos de datos jurídicos, nos invita a reflexionar sobre los nuevos retos de la informática jurídica documental, dado que, en la sociedad de la información actual, el éxito de cualquier tarea, negocio o trabajo, consistirá en explotar y optimizar la información en la mayor medida de lo posible, para reinar en la era del conocimiento.

Las consecuencias de no adoptar, los nuevos cambios del fenómeno del *Big Data*, son nefastas. La enorme cantidad de información, puede acarrear confusión para el operador jurídico, ya que si la base de datos jurídica es plana y no muestra la información, dentro de un contexto con sus pros y sus contras, podrá ocasionar confusión al usuario, quien puede tomar una decisión errónea o inconveniente, dentro de su actividad. De igual forma, la velocidad y el flujo de datos puede afectar la interpretación jurídica que se efectúe, sobre un tema en concreto, piense usted, en el caso en que la actualización de una base de datos no haya incluido, en tiempo real, la última jurisprudencia que tiene efectos de cosa juzgada y que puede afectar un litigio en curso.

Los bancos de datos jurídicos se alimentan de una enorme cantidad de documentos que necesitan ser captados, procesados, analizados en un tiempo record, otorgando

resultados e información relacionada y situada en un contexto de búsqueda particular. Podemos decir que las nuevas bases de datos jurídicos deberán ser oráculos del derecho, en donde, incluso, se conseguirá prever las tendencias jurídicas y la evolución del derecho, en un futuro. En el fenómeno del *Big Data*, lo relevante no es el gran conjunto de datos sino crear inteligencia artificial, a partir de esos datos.

3.2. El análisis de los datos jurídicos en la era del Big Data

El análisis de los documentos jurídicos, elaborado por abogados y personas especializadas en bancos de datos jurídicos y que se utiliza para integrar el contenido de dichos documentos dentro del tesoro jurídico, impide que la actualización de la información se realice en tiempo record y real.

Para limitar el tiempo de análisis, es necesario contar con la ayuda de herramientas de hardware y software, que analicen la información eliminando costos y tiempo. En este sentido, “las tecnologías del *Big Data* facilitan la toma de decisiones en tiempo real, comparando incluso con datos históricos”⁴⁹⁶.

Adicionalmente, como los datos contenidos en los bancos de datos jurídicos son datos estructurados, es posible que una máquina pueda gestionar la información, gracias a las tecnologías de análisis de texto y de procesamiento de información.

⁴⁹⁶ JOYANES AGUILAR, L. (2013). *Análisis de grandes volúmenes de datos en organizaciones*. México: Alfaomega Grupo Editor. p. 67.

3.3. La era de la web 3.0 o Web Semántica

El concepto de web semántica obedece a la necesidad inherente en la mutación masiva e innegable que tiene el uso de las nuevas tecnologías, en todos y cada uno de los entornos, en los cuales, el hombre como constante de la ecuación social, se ha desarrollado con el propósito fundamental de suplir sus necesidades y expectativas.

Ajustando el tema de este acápite, la necesidad a la cual debemos hacer mención, no es otra más que la de aplicar el potencial de la Web en el uso, manipulación y divulgación de información que, está por demás decirlo, llega a alcanzar niveles estratosféricos.

Esta es, sin duda, una de las realidades que más ha tomado desprevenidos a los usuarios y, en general, a todos los agentes que participan día a día en el tráfico mediático de información que se materializa en el ecosistema del internet, ya que, desde la creación del ARPANET⁴⁹⁷, en 1968, con su implementación exclusivamente académica y hasta la materialización de modernas plataformas de compras, medios de pagos, redes sociales y otros modelos de negocio en línea, nadie se imaginó el desmesurado crecimiento exponencial de los aplicativos virtuales y, mucho menos, de la cantidad de información depositada en medios digitales sistematizados.

⁴⁹⁷ Primera red gestora de comunicaciones entre usuarios y servidores de diferentes lugares de Norte América. Creada por encargo del Departamento de Defensa de Estados Unidos ("DOD" por sus siglas en inglés), a través de la *Agencia de Proyectos de Investigación Avanzada (ARPA)* como medio de comunicación para los diferentes organismos del país. Visto en línea en <http://es.wikipedia.org/wiki/ARPANET>

Según un estudio realizado por la revista *Science*, publicado en abril del 2011, para ese mismo año, la humanidad a lo largo de su historia, habría generado un equivalente a un trillón de bytes de información, de la cual, el 99% estaría soportada en formato digital y solo un reducido 1% reposaría en fuentes físicas, como el papel⁴⁹⁸. Lo anterior es solo un claro ejemplo de lo manifiesta que es la situación contemporánea de los datos en la web, de su magnitud y de su heterogeneidad que, valga la pena resaltar, sigue creciendo con el pasar del tiempo y en muy superiores proporciones, en comparación con los orígenes del internet.

Debe pensarse entonces que, obviando el interrogante del almacenamiento masivo de toda la información generada por las personas en su diario vivir, el acceso a estos datos debe estar garantizado y protegido, de tal suerte que su posterior consulta y manipulación, se realice de forma dinámica y sencilla, permitiendo que, los usuarios de los mismos, no encuentren obstáculos en sus intereses al navegar por la World Wide Web, fenómeno que no es ajeno a los juristas

Con todo, la categorización masiva de información y su vinculación mediante las diversas plataformas de búsqueda, en internet, no resulta una labor sencilla y cien por ciento productiva, pues, piénsese, es casi igual de engorroso sistematizar y evaluar cada dato como fabricar o generar el mismo.

En ello consiste la importancia de la web semántica, la cual busca que, por medio de la construcción de niveles de sucesiva abstracción y control, se realice un filtrado automático de datos estructurados, permitiendo posteriormente el intercambio de

⁴⁹⁸ MARTÍN, H., & PRISCILA, L. (2011). *The World's Technological Capacity to Store, Communicate, and Compute Information*. Science Magazine, 332(6025), p. 60.

estos mismos metadatos digitales, mediante la implementación de semántica, modelos lingüísticos, y abstracciones ontológicas. La recepción, estructuración y gestión general de los datos, están determinados por el procesamiento automático de los metadatos, por agentes inteligentes que sustituyan, en todo o en parte, la actividad propia del operador humano.

En consecuencia, La web 3.0 es el producto de la evolución de Internet. La interacción en la red y la enorme cantidad de información han abierto el camino para nuevas transformaciones. Internet es una enorme base de datos, ni los buscadores de información, verbigracia: Google, Yahoo, Ask, logran indexar la totalidad de la información que reposa en Internet, de ahí la existencia de la web invisible.

Como se ha mencionado, el surgimiento del *Big Data* o los grandes datos, hacen que la búsqueda de la información y la obtención de resultados pertinentes se tornen más difíciles. En tal sentido, nos encontramos en frente de la web 3.0 que surge en respuesta a la evolución de Internet. La Web 3.0 transforma el uso y la interacción entre el usuario e Internet.

Es menester señalar que “la Web Semántica fue creada y diseñada por el padre de la Web, Tim Berners-Lee, sin embargo ya su iniciativa ha pasado en estos dos últimos años de los cenáculos de las universidades y centros de investigación a las empresas y de ahí poco a poco al gran público”⁴⁹⁹.

⁴⁹⁹ JOYANES AGUILAR, L. (2009). *Redes sociales: la nueva revolución de la era de facebook*. España: Joxman Editores Multimedia. p. 44.

Dentro de esta nueva era, se han impulsado dos nuevas tecnologías, a saber: La web semántica y la inteligencia artificial. La tecnología de la Web 3.0 o web semántica surge como programa inteligente que parte de datos semánticos recopilados, con el fin de optimizar el uso de los mismos, obteniendo resultados eficientes en la utilización de grandes cantidades de información.

La web semántica es la web del significado, la web de los datos, que tiene por idea principal incluir el contenido, el significado y la relación de la información contenida en Internet, en un sitio web, para que una máquina de procesamiento pueda evaluarlas automáticamente. Lo anterior, con la finalidad de mejorar el Internet, mediante el uso de programas inteligentes que buscan y arrojan información oportuna, correspondiente con el tema objeto de la búsqueda.

El término Web semántica se utiliza, igualmente, para hacer referencia a la evolución de la inteligencia artificial, donde el sistema ya no solo se limita a almacenar, clasificar y buscar información sino a examinar su contexto, su significado y su pertinencia dentro del objetivo de la búsqueda.

Las páginas de la Web Semántica o la “web 3.0”⁵⁰⁰ contienen más información y significado que el que ofrece una página web 2.0 o 1.0. Las páginas Web 3.0 son etiquetadas y marcadas de manera tal que las máquinas puedan hacer sentido semántico de ellas, tal como los humanos hacemos sentido semántico del lenguaje.

⁵⁰⁰ HENDLER, J. (2009). *Web 3.0 emerging*. IEEE Journals & Magazines, 42(1). doi:10.1109/MC.2009.30. p. 89.

De esa forma, las máquinas pueden hacer conexiones instantáneas que tomarían mucho tiempo realizar o hasta reconocer, a las personas⁵⁰¹.

Lo anterior explica por qué la Web, por ser creación humana, se ha concebido como una herramienta para ser entendida por los seres humanos; de allí que la Web Semántica busca ir más allá y se enfoca en entender y manejar mejor la información que está en la Web y en permitir tareas más complejas a las máquinas⁵⁰².

La Web Semántica, propone superar las limitaciones de la Web actual, mediante la introducción de descripciones claras del significado, la estructura interna y la estructura global de los contenidos y servicios disponibles en la Web⁵⁰³. “Frente a la semántica implícita, el crecimiento caótico de recursos, y la ausencia de una organización clara de la Web actual, la Web Semántica intenta clasificar, dotar de estructura y anotar los recursos con semántica explícita procesable por máquinas”⁵⁰⁴.

A su vez, la Web Semántica ha sido definida por la organización World Wide Web Consortium (W3C) como:

⁵⁰¹ START MAKING SENSE:. (2008, Abril 9). *The Economist: Big and small companies are getting into the business of building an intelligent web of linked data*. Consultado en junio 23, 2015.

<http://www.economist.com/node/11002939>

⁵⁰² BENJAMINS, R., BREUKERS, J., CASANOVAS, & GANGEMI, A. (2005). *Law and the Semantic Web, an introduction*. *Law and the Semantic Web*, p. 1-7.

⁵⁰³ CASTELLS, P. (2003). *La Web Semántica*. In C. Bravo Santos, & Á. Redondo Duque, *Sistema interactivos y colaborativos en la Web*. Cuenca: Universidad de Castilla-La Mancha. p. 195.

<https://goo.gl/6CXQrO>

⁵⁰⁴ Ibídem. p. 200.

“una Web extendida y dotada de mayor significado, apoyada en lenguajes universales, que van a permitir que los usuarios puedan encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a una información mejor definida. Con esta Web, los usuarios podrán delegar tareas en el software que será capaz de procesar el contenido de la información, razonar con éste, combinarlo y realizar deducciones lógicas para resolver automáticamente problemas cotidianos”⁵⁰⁵.

La Web Semántica es así una Web de información, de fechas, de títulos, de números, de propiedades y cualquier otro dato o información que pueda ser concebido.

Es importante mencionar que el objetivo principal de las tecnologías de la Web Semántica, es desarrollar una Web más cohesionada, donde la información sea aún más fácil localizar, compartir e integrar, para de esta forma sacar mayor provecho de los recursos disponibles en la Web⁵⁰⁶.

La iniciativa de la Web Semántica tiene como objetivo adicional definir el lenguaje y las estrategias para dotar a la Web tradicional de representaciones del conocimiento procesables por máquinas. “La Web Semántica es un esfuerzo de crear un marco de representación y organización de la información contenida en mensajes y compartida en la Internet ese enorme repositorio mundial al que toda la información disponible se ve abocada, a partir de sus relaciones profundas en el plano del significado. De ahí viene el uso del adjetivo semántico. Y ese énfasis en lo

⁵⁰⁵ WORLD WIDE WEB CONSORTIUM. (n.d.). Guía breve de Web Semántica. Consultada en noviembre 8, 2014. <http://www.w3c.es/Divulgacion/Guiasbreves>.

⁵⁰⁶ CASTELLS, P. (2003). Op. cit. p. 200.

semántico nos catapulta sin remedio a la comprensión del mundo del conocimiento, esas estructuras que están subyacentes a los mensajes en nuestro caso a los documentos, y más específicamente en su forma de recursos Web y que han sido lentamente iluminadas por filósofos y lingüistas”⁵⁰⁷.

Ahora bien, el impacto y la importancia de la Web Semántica ha sido resaltado por el autor Francisco Javier García Marco, quien señala que “la Web Semántica es uno de los dos nuevos horizontes de la World Wide Web, el conjunto de tecnologías que, sobre la Internet, ha revolucionado la forma en que publicamos, almacenamos, recuperamos e intercambiamos la información. El otro es, sin duda, la Web Social o Web 2.0. Las dos son evoluciones lógicas de la Web concebida en 1989. La primera la Web Semántica supone más y mejor computación distribuida entre los ordenadores que forman la Red; la segunda, más y mejor comunicación entre las personas que usan la Red. Detrás de ambas, la satisfacción del usuario, que gana en potencia de actuación, autonomía, seguridad, relación social, eficacia y eficiencia, aunque, ciertamente, no se deben obviar, por otro lado, los muchos peligros a conjurar, relacionados con la calidad y autenticidad de la información, su preservación y sus implicaciones para la privacidad.”⁵⁰⁸

La afirmación de este autor converge en que la Web Semántica es una herramienta que hace mucho más eficiente el proceso de búsqueda de información y de igual manera su surgimiento beneficia a los usuarios en la organización de los contenidos y la relevancia de los mismos.

⁵⁰⁷ GARCÍA MARCO, F. J. (2009). *Perspectivas sobre el uso de la Web Semántica en el tratamiento de la información y documentación legislativa*. Revista Scire, 15(1), p. 59.
<http://ibersid.eu/ojs/index.php/scire/article/view/1769/1740>

⁵⁰⁸ *Ibíd.*

Entre otras ventajas de la Web Semántica, encontramos las siguientes:

1. “La Web semántica ayuda a las máquinas a entender las fuentes que se encuentran en la Web en donde la información puede estar compartida y procesada por herramientas automatizadas, como máquinas de búsqueda y usuarios humanos.
2. La Web semántica mantiene los principios de descentralización, compartición, compatibilidad, máxima facilidad de acceso y contribución. En este contexto un problema clave es alcanzar un entendimiento entre las partes que han de intervenir en la construcción y explotación de la Web como son: usuarios, desarrolladores y programas de diverso perfil. La Web semántica rescata la noción de ontología del campo de la Inteligencia Artificial como medio para cumplir este objetivo”⁵⁰⁹.
3. La Web Semántica es una tecnología emergente que promete incrementar la habilidad para que la computadora pueda analizar la información, dando como resultado, aplicaciones más inteligentes, motores de búsqueda más eficientes y mejoras a nuestra habilidad para recopilar y procesar la información. Aplicado al Derecho, la Web Semántica puede tener un efecto transformador en la forma en que los abogados manejamos la información, toda vez que esta se constituye como el elemento clave, en el ejercicio de la profesión.

⁵⁰⁹ BAADER, F., HORROCKS, I., & SATTler, U. (2015, Febrero 2). *Description Logics as Ontology Languages for the Semantic Web*. *Mechanizing Mathematical Reasoning*. 228-248. Consultada en febrero 2, 2015. <http://www.sfu.ca/~mhatala/iat881/papers/baader-et-al.pdf>

4. La Web Semántica traerá estructura y nuevos significados a las páginas web, creando un ambiente donde el software puede realizar tareas muy sofisticadas para los usuarios. La Web Semántica no es una Web separada sino una extensión de la actual, en la cual a la información se le provee un significado bien definido, permitiendo a los usuarios y computadores, trabajar en cooperación. Con los esfuerzos actuales, a través de la Web Semántica, las máquinas serán mucho mejores en el procesamiento y comprensión de los datos e información que actualmente se presentan a los usuarios⁵¹⁰.

La aplicación de la web semántica, en la informática jurídica, reviste un total interés porque Internet hoy en día permite visualizar gracias a HTML sitios web jurídicos y no jurídicos. El código HTML puede precisar que el tema de un documento es por ejemplo “la prescripción”, con la web semántica sería incluso posible precisar dentro del código HTML aspectos específicos, verbigracia, si el documento se trata de la prescripción adquisitiva o de la prescripción extintiva, o si es otro tipo de prescripción. De igual forma con la aplicación de la web semántica en la informática jurídica, sería posible en la actualidad, relacionar u oponer entre si los datos contenidos en una o varias bases de datos.

La web semántica aparece, entonces, como una respuesta para buscar, restringir y filtrar términos, conceptos, relaciones, metadatos y meta tesauros⁵¹¹, tal y como se

⁵¹⁰ BERNERS-LEE, T., HENDLER, J., & LASILLA, O. (2001). *The Semantic Web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities*. Scientific American, 284(5), p. 34.

⁵¹¹ El metatesauro es un tesauro enriquecido, que contiene información semántica y lingüística, adicional

constituye una red semántica. Ésta proporciona los elementos básicos que se podrían emplear para dar solución al problema del significado de los conceptos presentados en la Web. Su gran funcionalidad y aplicación se extiende a diversos campos del conocimiento, entre ellos al Derecho, un campo que presenta gran complejidad en la recopilación y manejo de su información⁵¹².

3.3.1. Web Semántica y su aplicación en derecho

La Web Semántica es una herramienta útil para la estructuración de la información, toda vez que le otorga un significado preciso y procesable de forma automatizada. Es claro que el beneficio más inmediato de las tecnologías semánticas está en poder organizar grandes cantidades de información de un área en particular, para hacerla sencilla de obtener y analizar. La aplicación de los principios de la Web Semántica al Derecho es un campo con enormes retos y potencial, si se tiene en cuenta la necesidad que tienen los abogados de hacer uso de la información jurídica, optimizando los tiempos de análisis de los documentos.

La Web Semántica en el campo del Derecho se puede entender como “un esfuerzo sistemático para organizar el conocimiento de la realidad social y de las técnicas jurídicas que se utilizan para estructurarla dentro de marcos políticos, tal y como aparecen de forma subyacente en la propia legislación, de manera que ese enorme

⁵¹² TORRES ANDRADE, C. A. (2008). *Agentes y web semántica: elementos claves para el futuro del e-learning*. Revista Tecnura, 11(22), p. 111.

conjunto de documentos sean capaces de responder de forma cada vez más sencilla a preguntas cada vez más inteligentes y sofisticadas”⁵¹³.

Es decir que, una aproximación de la Web Semántica, aplicada al Derecho, va más allá de la mera gestión y organización de la información jurídica, para plantear un sistema de gestión y manejo de la información legal eficiente, coherente y unificado que permita relacionar, jerarquizar y equiparar conceptos jurídicos para que los datos jurídicos en su conjunto, hablen por sí solos y se obtengan resultados semánticos.

Los sistemas de información jurídicos, en su mayoría, incorporan distintos modelos de herramientas terminológicas, escasamente estructuradas en el sentido documental, mostrando una clara preferencia por las listas de palabras clave, listas autorizadas de descriptores y tesauros⁵¹⁴.

Las tendencias más recientes se dirigen a soluciones automatizadas de expansión semántica y sugerencias automáticas, en las que se tiene en cuenta la relevancia, atendiendo tanto a los contenidos como al rango de la disposición normativa o del tribunal, en el caso de las bases de datos de jurisprudencia⁵¹⁵.

Es muy difícil poder medir y predecir el impacto de la Web Semántica en el Derecho y en la cultura en general, ya que son muy pocos los estudios existentes en la actualidad. Lo que sí es posible afirmar es que se crearán nuevos contextos y

⁵¹³ GARCÍA MARCO, F. J. OP. CIT.

⁵¹⁴ CODINA, L., & PEDRAZA JIMÉNEZ, R. (2011, Septiembre- Octubre). *Thesaurios y ontologías en sistemas de información documental*. El profesional de la información, 20(5), p. 555.
<http://www.elprofesionaldelainformacion.com/contenidos/2011/septiembre/10.pdf>

⁵¹⁵ ALVITE DÍEZ, M. L. (2012). *El uso de vocabularios controlados en sistemas de información jurídica: evolución y tendencias actuales de representación*. Revista Scire, 18(1). p. 29.

ambientes de interacciones, para los abogados como parte del ejercicio de la profesión. Es un campo con muchos retos interesantes que, sin duda alguna, tendrá un impacto importante en el mundo jurídico.

El avance de la tecnología está llevando al Derecho, a experimentar cambios muy rápidos en términos de redacción de documentos, proyectos legislativos y gobierno en línea. Se puede anticipar que la Web Semántica tendrá un importante impacto en la evolución de la documentación, ya que es probable que la funcionalidad avanzada permita revisar la consistencia interna de un documento o el cumplimiento de una serie de leyes o regulaciones específicas⁵¹⁶.

El autor Casanovas Pompea nos ofrece algunas aproximaciones jurídicas a la Web Semántica, a saber:

- **“Akoma Ntoso:** Akoma Ntoso (“corazones unidos”), constituye un conjunto de estándares en XML automáticamente legibles para el intercambio y reusabilidad de documentos judiciales, legislativos y parlamentarios. Es decir, se sitúa al nivel descriptivo de los distintos niveles de los documentos oficiales para su clasificación, búsqueda y uso posteriores”⁵¹⁷

⁵¹⁶ HERNÁNDEZ MARIN, R. (2009). *Sobre ontología jurídica e interpretación del derecho*. Alicante: Biblioteca Miguel de Cervantes. p. 34.

⁵¹⁷ CASANOVAS, P. (2012, Enero-Junio). *Algunas líneas de investigación en gestión del conocimiento jurídico: web semántica, ODR y derecho racional*. Revista Scire, 18(1), p. 18.
<http://ibersid.eu/ojs/index.php/scire/article/viewFile/3943/3672>

Es “Akoma Ntoso”⁵¹⁸, una de las plataformas internacionales más destacadas en la implementación de tecnologías de la información como recurso primigenio en el estudio, análisis y comunicación de documentos jurídicos. Se ha convertido en un estándar mundial para las bases de datos jurídicas que funcionan on line y que tienen por propósito, la democratización de la información jurídica, no como simples archivos de texto sino como archivos estructurados, fáciles de consultar, reproducir e interpretar, siempre bajo la premisa fundamental de que cada dato representa una vertiente más, en ese todo que puede ser denominado ordenamiento jurídico.

Dentro de los objetivos estratégicos, establecidos en la página web del Akoma Ntoso, se enumeran:

- ✓ Crear una lengua franca, natural o neutral que permita gestionar la interrelación de las diferentes naciones africanas que se empeñan en crear un estándar documental jurídico amplio y suficiente.
- ✓ Proporcionar un entorno de almacenamiento, a largo plazo, para contener toda la información y documentos parlamentarios, legislativos y judiciales y que facilite su consulta.
- ✓ Facilitar el proceso de producción normativa y jurisprudencial, en las diferentes instituciones y corporaciones. Reducir la multiplicidad de proyectos autónomos que persiguen iniciativas de TI en el área de la

⁵¹⁸ UNIVERSITY OF BOLOGNA. (2015). *Akoma Ntoso: XML for parliamentary, legislative and judiciary documents*. <http://www.akomantoso.org/>

producción parlamentaria y la gestión de documentos legislativos y judiciales.

- ✓ Crear un conjunto de datos comunes y modelos de metadatos, de manera que las herramientas y técnicas de recuperación de información que se utilizan en un Parlamento o Corte determinando, también pueden ser utilizadas en diferentes Parlamentos o Cortes, de manera coherente y eficaz.
 - ✓ Crear mecanismos comunes para la asignación de nombres de recursos y la vinculación, de modo que los documentos producidos por los Parlamentos / Tribunales, puedan ser fácilmente citados y referencias cruzadas - ya sea por otros parlamentos / Tribunales o por otros usuarios.
 - ✓ Construir una herramienta de fácil utilización y “auto-explicativa” al proporcionar por si sola toda la información para su uso y significado, a través de un simple examen, incluso sin la ayuda de software especializado.
 - ✓ Ser "extensible", es decir, debe ser posible permitir modificaciones a los modelos en el marco Akoma Ntoso, de manera que se pueda dar una personalización local que garantice la interoperabilidad con otros sistemas.
- **MetaLex** busca esencialmente lo mismo que Akoma Ntoso. Uno empezó

como un programa para el análisis y clasificación de los documentos en los Parlamentos africanos. El segundo, que es anterior al primero, como un conjunto de estándares para la descripción de las leyes holandesas. Ambos conjuntos coincidieron en el proyecto europeo ESTRELLA (2007-2009)⁵¹⁹ y recogieron experiencias previas como la danesa LexDania y la italiana Norma-in-Rete.⁵²⁰

- **El enfoque del IDT:** “el IDT propone una perspectiva conceptual complementaria, centrada en los datos y situaciones complejas de uso socialmente relevantes. (...) Así, los elementos de negociación, resolución y gestión de conflictos entran también como un objeto a tener en consideración en las redes sociales, en el comportamiento del usuario, y en la multiplicidad de actos que éste puede realizar a través de la red⁵²¹”.

David Siegel ve un enorme poder de la Web Semántica al transformar el trabajo de los abogados. El considera que las taxonomías legales y reglas formalizadas pueden resultar en “un conjunto de reglas semánticas que pueden servir como una versión del derecho legible por las máquinas”⁵²². Este es un punto muy interesante que podría generar debates sobre la futura utilidad de los abogados en un mundo en el cual, mucho de lo que hacen, puede ser realizado por máquinas. Sin embargo, pensamos que es posible que las nuevas herramientas y la aplicación de la Web Semántica al Derecho, vayan a facilitar las tareas legales básicas del derecho, sobre

⁵¹⁹ Navegar el Sitio Web del Proyecto en <http://www.estrellaproject.org/>

⁵²⁰ CASANOVAS, P. (2012).

⁵²¹ Ibídem.

⁵²² Siegel, D. (2009). The power of the Semantic Web to transform your business group. Penguin Group. Simple HTML Ontology Extensions. SHOE. (n.d.). p. 187.
<http://www.cs.umd.edu/projects/plus/SHOE/>

todo en la búsqueda de información, pero difícilmente puedan en el corto plazo superar el análisis más complejo y profundo que puede hacer un abogado de un caso o situación particular⁵²³.

La estandarización de los documentos legales y la interoperabilidad institucional en cuanto a su metodología, conceptualización y desarrollo, permiten hablar no solo de documentos jurídicos estructurados, sino también, de metadatos inclusivos, que son de gran relevancia para el análisis del documento: idioma, medio de publicación, procedimiento de expedición, normas concordantes, etc. Estas categorías, como características adicionales al contenido mismo del documento, facilitan su interpretación, de tal suerte que los diferentes agentes jurídicos, que no necesariamente tienen que ser juristas o abogados calificados, vean superadas las barreras lingüísticas propias de la disciplina jurídica.

Este escenario, que valga decir tiene alcance más allá del derecho, se concreta, de forma específica bajo el concepto de “*interoperabilidad semántica*”, el cual, se refiere a que cualquier documento legal (por ejemplo los debates parlamentarios, leyes, decretos, sentencias, etc.) que goce de una estructura de datos, con un nivel semántico mínimo, generaran una determinada sintaxis que permitirá representar dicha estructura en un formato comunicable, por vía informática y con independencia del sistema o plataforma receptora. Para que esta interoperabilidad garantice la búsqueda e interpretación efectiva de la información comunicada a los usuarios de la misma, es necesario que se implementen, además de los modelos de

⁵²³ HARLEY, B. (2010, Marzo 31). *Semantic Lawyering: How the Semantic Web Will Transform the Practice of Law* (Part 1). Columbia Science and Technology Law Review, Parte 1. Semantic Lawyering: How the Semantic Web Will Transform the Practice of Law: <http://stlr.org/2010/03/31/semantic-lawyering-how-the-semantic-web-will-transform-the-practice-of-law-part-1/>

metadatos, herramientas que faciliten la representación formalizada del contenido de la información: *lenguajes de indexación*⁵²⁴.

Dicho esto, piénsese en lo importante y necesario que es diseñar un arquetipo semántico para la web, máxime, cuando se tienen propósitos afines a los enunciados y establecidos para el Akoma Ntoso. Vincular el conocimiento legal y su expresión documental, mediante la teoría jurídica y la perspectiva que esta tiene sobre el entorno web, amplía el margen regulatorio de un determinado ordenamiento jurídico. El manejo de la información, su archivo, almacenamiento y estructuración para posteriores consultas en la web, cobran un mayor y más claro matiz cuando se pone de presente el potencial social, que lleva implícito, dicha gestión, en el marco de lo que podría denominarse una “eficiente administración de justicia”.

En concordancia, la web semántica aplicada al derecho, se presenta como útil y pertinente por sí misma, ya que, con el agregado de lograr recopilar toda la información legislativa, jurisprudencial y doctrinaria, existente y de forma digital, integra patrones de búsqueda eficaces, en los que se atiende no solo a conceptos técnicos o eminentemente jurídicos.

La web semántica, proporciona a sus usuarios arquetipos de búsqueda en los que, con las taxonomías suficientes, la exploración de un resultado determinado, se sujeta a criterios que pueden ser proporcionados tanto por profesionales en el derecho (lenguaje especializado), como por un ciudadano del común (lenguaje natural).

⁵²⁴ ALVITE DÍEZ, M. L. (2012). op. cit. p. 30.

En conclusión, la Web Semántica tiene un importante impulso que dilucida el próximo paso en la evolución de la Web. La dirección real de ese impulso, presenta un gran reto tanto desde el punto de vista regulatorio como del impacto y la aplicación que la Web Semántica tendrá en el Derecho. El superar los retos de la Web Semántica va a depender de la habilidad con la cual los miembros de la profesión legal nos aboquemos a su estudio y preparemos soluciones semánticas a los retos que se vayan presentando.

3.3.2. Herramientas de la Web Semántica

El avance de la Web Semántica se ha basado en el desarrollo de importantes tecnologías. Entre ellas podemos mencionar varios componentes: Extensible Markup Language (XML), el Resource Description Framework (RDF) y el Web Ontology Language (OWL). Algunas de estas tecnologías existen, incluso, desde la web 1.0 y las otras forman parte de la web 2.0.

El primer lenguaje para la construcción de la Web Semántica fue SHOE, creado por Jim Hendler en la Universidad de Maryland en 1997⁵²⁵. SHOE es una pequeña extensión de HTML que permite a los autores de páginas Web anotar o marcar sus documentos Web, con conocimiento procesable por las máquinas. A partir de SHOE han surgido importantes propuestas entre las que se destacan las siguientes:

⁵²⁵ *Simple HTML Ontology Extensions. SHOE*. (n.d.). Consultada en junio 23, 2015.
<http://www.cs.umd.edu/projects/plus/SHOE/>

EL XML que permite a todos los usuarios crear sus propias etiquetas, es decir, marcas ocultas que realizan anotaciones en secciones del texto de una página Web. Los scripts o programas pueden hacer uso de estas etiquetas en forma muy sofisticada. En pocas palabras el XML admite a los usuarios incluir estructura arbitraria a sus documentos, pero no dice nada sobre que significa dicha estructura⁵²⁶, cabe subrayar que este lenguaje existe desde la Web 2.0 y que ya ha sido objeto de análisis en esta tesis, en el capítulo anterior.

Por su parte, el RDF es un lenguaje, basado en XML, dotado con una semántica para “hablar de la Web”. “El RDF asume que existe un mundo en el que existen recursos y relaciones entre recursos, y el lenguaje se orienta para que un usuario declare hechos que involucren esas relaciones y esos recursos”⁵²⁷. Esta herramienta permite estandarizar las palabras claves y facilitar el procesamiento en las máquinas.

El OWL aporta el vocabulario estándar para representar ontologías, las herramientas conceptuales más potentes de las que disponemos para expresar semántica al más alto nivel. Las ontologías, como se explicará con mayor detalle más adelante en este trabajo, han adquirido un papel relevante y proliferan en todo tipo de sistemas de información, como herramientas que soportan las búsquedas conceptuales y otra serie de funcionalidades⁵²⁸. El OWL recoge las palabras claves estandarizadas por el RDF y les da sentido, es uno de esos lenguajes que, además, está constituido por una forma particular de lógica llamada lógica descriptiva. Las lógicas descriptivas son

⁵²⁶ BRAY, T., & BOSAK, J. (1999). *XML and the second-generation web*. Scientific American, 280(5), p. 89-93. <https://courses.washington.edu/infx598/win12/secondGenerationWeb.pdf>

⁵²⁷ DÁVILA, J. (2006). *La lógica, los agentes y la web semántica*. Revista de la Universidad de los Andes de Mérida. p 25.

⁵²⁸ ALVITE DÍEZ, M. L. (2010). op. cit. p. 47.

“un desarrollo reciente dirigido a lenguajes que se prestasen para el tratamiento computacional, incluso si eso implicaba restringirlos para que ciertas declaraciones (problemáticas por su complejidad computacional) no se pudieran decir en ellos⁵²⁹”.

Hay otros desarrollos intermedios entre RDF y OWL, como RDF-Scheme y DAML-OIL. No se trata de extensiones solamente, sino de diversos lenguajes cubriendo diversos aspectos del problema.

El uso de la Web semántica y la interconexión de recursos puede optimizarse etiquetando documentos con información semántica, de manera que la información es interpretada por un lector automático, que realiza la actividad de un humano. Estas etiquetas pueden contener metadatos que describan aspectos suplementarios de los documentos.

3.3.3. Metadatos en la Web Semántica - importancia en la informática jurídica

En este punto es preciso enervar un término que sería acuñado muchos años antes de que se entrara en conciencia sobre la multiplicidad informativa en el ciberespacio. El *metadato*, término acuñado por Jack Myers en la década de los 60⁵³⁰, y que es entendido como el conjunto de datos o sistema estructurado de datos, que es utilizado para describir otro dato o conjuntos de datos y que, por demás, sirve para catalogar información, sirviéndose de un sistema determinado que permite identificar y describir el *dato* según su contexto, calidad y condición.

⁵²⁹ *Ibidem*.

⁵³⁰ CAPLAN, P. (1995). *You Call It Corn, We Call It Syntax-Independent Metadata for Document-Like Objects*. The Public-Access Computer Systems Review, 6(4).
[//journals.tdl.org/pacsr/index.php/pacsr/article/view/5992/5621](http://journals.tdl.org/pacsr/index.php/pacsr/article/view/5992/5621)

Ahora, la naturaleza de los metadatos y la reproducción de gestores de búsqueda abiertos dedicados no solo a la localización sino también al análisis de la información, permite que la interoperabilidad de los sistemas y plataformas web se vea en gran medida beneficiada por cuanto, la redistribución y reutilización de datos y metadatos en una “*infraestructura común*” implicaría por sí misma un ahorro de recursos para los operadores de redes de información, así como un aumento en la fluidez de usuarios de estas redes, por la facilidad de acceso, ejecución y reproducción de los recursos que aún continúan disponibles.

Los metadatos son herramientas de la Web semántica e igualmente constituyen una herramienta útil dentro del campo de la informática jurídica, toda vez que gracias a su existencia, es posible crear nuevos datos a partir de los documentos jurídicos recopilados.

El significado de la palabra metadatos, quiere decir, más allá del dato o sobre datos. Los metadatos, tienen por finalidad realizar descripciones de otros datos y permitir su localización o ubicación dentro de un conjunto de datos. Ellos pueden presentarse en forma física o digitalizada.

Para la informática jurídica, los metadatos se constituyen como un componente esencial a la hora de recuperar y organizar la información. En la Web semántica, igualmente, los metadatos de cada documento permiten realizar enlaces relacionales entre un documento y el otro. Estos metadatos, admiten filtrar y restringir la información de forma que el resultado se limite al verdadero objeto de la búsqueda.

Mejorar los criterios de búsqueda implica como es obvio, mejorar la búsqueda en sí. Agrupar la información y clasificarla de forma descriptiva a través de los metadatos, lleva necesariamente a diseñar todo un sistema de catalogación, que integre y desarrolle estrategias y mecanismos para la búsqueda y uso de la información (servicios)⁵³¹.

De forma paralela, los metadatos traen consigo, muchos otros beneficios: la proliferación de la información en sus versiones digitales por medio de Internet; la precisión en las búsquedas mediante la implementación de descriptores y restrictores; la disminución masiva del uso de ancho de banda en todo el mundo, al ser más liviano el *metadato* respecto del *dato*, en cuanto a “bits” se refiere; el consolidar derechos como el de autor y en general, el de la propiedad intelectual, al permitir un mejor control del acceso y uso de la información, según parámetros de licenciamiento y permisos.

En la informática jurídica, los metadatos categorizan los documentos jurídicos mostrando la eventual relación entre los mismos. Este mismo método se utiliza cuando se pretende implementar un sistema de inteligencia artificial aplicado al derecho, con el fin de obtener conclusiones automáticas.

Para efectos de eficacia de los metadatos dentro de la informática jurídica, se hará necesario entre otras cosas, las siguientes:

⁵³¹ AGUDELO BENJUMEA, M. M. (s.f). *Los Metadatos. Gestión de Contenidos de Educación Virtual de Calidad*. Ministerio de Educación Nacional de Colombia y Universidad de Antioquia. Consultada en febrero 12, 2015. http://aprendeenlinea.udea.edu.co/lms/men/docsoac3/0301_metadatos.pdf

- ✓ Que los metadatos realicen una descripción de cada documento jurídico y en particular de su estructura. En esta primera etapa de creación, es posible trabajar de forma automatizada, semiautomatizada o manual. El volumen de datos, así como la complejidad de los términos y contenidos de los documentos, definirá cual solución se adapta en el caso concreto para crear los metadatos.

Es importante resaltar que, para efectos de la informática jurídica, bien podría adaptarse un proceso semiautomático con el fin de crear los metadatos, toda vez que es poco probable que un software pueda extraer todos los metadatos de forma autónoma. Afirmar lo contrario, sería adoptar una posición irreal, ya que el apoyo de un jurista especializado en la materia a analizar, constituye una herramienta que garantiza la disminución considerable de inseguridades, errores e imprecisiones en la terminología jurídica, tarea que un software especializado en extracción de metadatos, no podría hacer de manera autónoma.

- ✓ Que se subdividan los metadatos semánticos que describen el contenido del documento de los que describen la estructura misma de su contenido.
- ✓ Que sean metadatos actualizados, si el contenido de una base de datos jurídica cambia sus metadatos tendrán que cambiar también, de manera que los metadatos, deberán ser adaptados a las nuevas realidades incluyendo las modificaciones, que han surgido dentro del ordenamiento jurídico. Existen metadatos jurídicos en los que las modificaciones se realizan de forma automática, gracias al uso de un software especializado, sin embargo, en la gran mayoría de casos, los bancos de datos jurídicos continúan siendo manejados por

operadores jurídicos humanos que modifican manualmente los metadatos.

- ✓ Se realice un control de actualización constante para generar su destrucción una vez el documento pierda vigencia, siempre dejando claro el motivo de su supresión, por ejemplo: derogación, subrogación, sustitución etc.

Habría que determinar en cada base de datos jurídica si resulta conveniente eliminar el metadato y su documento. La razonabilidad es el criterio que define si se conserva o no un metadato. Esto queda a discrecionalidad de quien dirige la base de datos.

A modo de ejemplo, si emprendemos la búsqueda en una base de datos jurídica sobre el “mandato”, este dato tendrá sus correspondientes metadatos clave adjuntos como: aceptación, mandato aparente, mandato tácito, ratificación, agencia oficiosa, responsabilidad del mandatario, revocación, obligaciones del mandante, obligaciones del mandatario, etc.

3.4. Importancia del uso de vocabulario semántico en los metadatos: las ontologías

Para garantizar la coincidencia, relación y afinidad de los metadatos, es importante fijar los términos abarcando los sinónimos, por ejemplo, en el caso de realizar la búsqueda de “agente inmobiliario”, la base de datos jurídica deberá localizar igualmente el mandato, o en el caso de introducir en la búsqueda la palabra “cónsul”, la base de datos deberá localizar igualmente todo lo relacionado con los agentes diplomáticos y consulares. Dentro de una base de datos jurídica, los metadatos

cumplen un papel fundamental, toda vez que asocian los documentos y sus palabras clave, de manera que se impida la pretermisión de información relacionada, que puede ser igualmente de utilidad.

En este punto las ontologías que definen justamente la relación entre los diferentes términos, cumplen un papel fundamental, porque su asociación permite que el software que elabora la búsqueda dentro de la base de datos, asocie los términos y arroje resultados mucho más completos. Por ejemplo, si el usuario de la base de datos jurídica ingresa en su búsqueda “promesa unilateral de venta”, es posible que la base de datos también arroje resultados sobre la ejecución forzada. El buscador, gracias al uso de las ontologías, comprende que la ejecución forzada forma parte de la promesa unilateral de venta, lo que ofrece una visión más completa del tema.

En este caso, contrario a lo que sucede con las folksonomías⁵³², que son definidas por los usuarios, piensen ustedes en la web 2.0 que es participativa, las ontologías son definidas por expertos jurídicos que clasifican y ordenan los términos generando asociaciones o relaciones entre los mismos.

Como se puede evidenciar, el uso de los metadatos y de las ontologías dentro de la informática jurídica, juegan un papel fundamental para facilitar la búsqueda y en especial para seleccionar la información pertinente, dentro del gran campo de información de una base de datos.

Si a estos conceptos de metadatos y ontologías le añadimos la aplicación de la

⁵³² La folksonomía, se entiende como aquella que se utiliza en la web social para describir el contenido del documento, mediante el uso de términos del lenguaje natural

lógica, las bases de datos jurídicas podrían empezar a razonar basadas en la inteligencia artificial.

3.5. ¿Cómo se genera inteligencia a partir de los datos jurídicos?

En el mundo del derecho el exceso de documentación jurídica, dificulta la búsqueda de un resultado adecuado que se adapte al contexto. El objetivo de esta tesis, como bien se ha mencionado, consiste en trasladar el uso de las tecnologías de inteligencia semántica al derecho, para que los usuarios al realizar sus búsquedas logren optimizar los resultados y aprovechar en una mayor medida cada uno de los documentos que reposa dentro de una base de datos jurídica.

La investigación sobre la Web semántica en la presente tesis, busca igualmente crear un sistema capaz de *razonar jurídicamente* a partir de una base de datos jurídica, basada en descripciones lógicas. Para tales efectos, la lógica juega un papel preponderante, toda vez que se hace necesario utilizar reglas en las cuales se exprese un razonamiento lógico entre los documentos y especialmente entre sus conceptos dentro de la base de datos.

En este sentido, resulta indispensable conocer ¿cómo se puede generar inteligencia a partir de los datos jurídicos? Tal vez, la pregunta precisa sería ¿cómo lograr que los datos que reposan en una base de datos jurídica hablen por si mismos?

3.5.1. Inteligencia artificial

La idea de simular la inteligencia humana no es nueva, ya en la época de la Segunda Guerra Mundial, se tenía una representación práctica del concepto, cuando Alan Turing, matemático, criptógrafo, filósofo, lógico y científico de la computación inglés, creó la máquina “bombe”, para descifrar los mensajes de *Enigma*, máquina creada por los alemanes, para cifrar sus comunicaciones, la ubicación y los ataques que el ejército alemán pretendía perpetuar.

Si el surgimiento de Internet se da gracias al programa de comunicaciones de las fuerzas militares estadounidenses, la inteligencia artificial tiene su origen en la necesidad de descifrar los mensajes emitidos por los alemanes durante el periodo que inicia en 1938. La máquina “bombe” contribuyó, durante este periodo de la historia, a anticipar los ataques y las intervenciones en un gran porcentaje de los casos; algunos, incluso aseguran que gracias a su creación el periodo de la guerra se pudo disminuir dos años.

Esta máquina “bombe”, reemplazó la actividad infructuosa de cálculo manual que realizaba un grupo selecto de científicos en Inglaterra, para descifrar el contenido de los mensajes de la máquina Enigma, toda vez que cada día de trabajo los códigos variaban y los científicos debían reiniciar el desciframiento de las comunicaciones, lo que dificultaba su labor, teniendo en cuenta además que existía un número de probabilidades infinitas que por circunstancias de tiempo era imposible llevar a buen término en tan sólo 24 horas. Esta máquina contaba con la capacidad de realizar

deducciones lógicas fundadas sobre las palabras probables, excluyendo las combinaciones incorrectas y arrojando sólo los resultados lógicos.

A su vez, esta máquina vino a ser la base de la computación existente actualmente. En el artículo “computing machinery and intelligence”⁵³³, se explica la posibilidad de crear inteligencia artificial y propone considerar la pregunta: ¿las máquinas pueden pensar? Este problema es descrito en términos de un juego que el autor denomina “el juego de imitación” o “*imitation game*”, en el juego participa un hombre (A), una mujer (B) y un interrogador (C), que puede ser de cualquiera de los dos sexos. EL interrogador se queda en un salón separado de A y B. El objetivo del juego consiste en que el interrogador determine cuál de los dos es el hombre y cual la mujer.

El interrogador C conoce a A y B, por etiquetas mujer y hombre, X y Y, de manera que al final del juego el podrá determinar que X es A y Y es B o que X es B y Y es A. El interrogador C puede formular preguntas tanto a A como a B. Por ejemplo:

C: ¿Podría X decirme la longitud de su cabello?

Suponiendo que X es en realidad A, A podría responder. A tiene como objetivo, inducir en error de identificación a C. Entonces su respuesta podría ser:

A: “Mi pelo es corto y las hebras más largas son cercanas a nueve pulgadas de largo”.

⁵³³ TURING, A. M. (1950). *Computing Machinery and Intelligence*. Mind, 49, 433-460.
<http://www.csee.umbc.edu/courses/471/papers/turing.pdf>

Con el fin de que los tonos de voz no ayuden al interrogador, las respuestas pueden ser escritas, o mejor aún, a máquina. La disposición ideal es contar con un teletipo de comunicación entre las dos habitaciones. De manera alternativa, las preguntas y las respuestas pueden ser repetidas, por un intermediario. El objetivo del juego, para el jugador B, es ayudar al interrogador. La mejor estrategia para ella es probablemente dar respuestas veraces. Ella podrá agregar frases a sus respuestas como “yo soy la mujer, no lo escuches a él”, pero es posible que el hombre realice comentarios similares.

Ahora nos preguntamos: ¿Qué pasaría cuando una máquina toma la parte de A, en este juego? ¿Podría el interrogador decidir erróneamente, a menudo, cuando el juego se juega de esta manera, como cuando el juego se realiza entre un hombre y una mujer? Estas dos preguntas rempazan la pregunta inicial ¿las máquinas pueden pensar?⁵³⁴

Una pregunta similar se podría formular en el campo de la informática jurídica: ¿es posible que una máquina pueda pensar, como lo haría un abogado o un juez? Si bien el razonamiento jurídico no ha sido desplazado por las máquinas, hasta ahora, la inteligencia artificial constituye una herramienta clave para ayudar a formar y robustecer el razonamiento jurídico. Las tendencias de su implementación han aumentado y hoy en día existen proyectos para materializar este objetivo.

⁵³⁴ *Ibíd.*

En este punto, cabe realizar una crítica a la expresión “inteligencia artificial”, toda vez que la inteligencia, desde el punto de vista humano, puede ser sometida a una evaluación, desde los diferentes campos del conocimiento, lo que no sucede con la inteligencia artificial, porque existen máquinas que actualmente pueden recrear tan sólo un sector o una parte del razonamiento humano. Verbigracia: *deep blue*, el computador especializado en el juego de ajedrez que, en 1997, logró ganar al campeón del mundo de ajedrez Garry Kasparov. *Deep blue* se encuentra especializado únicamente en ajedrez, por lo cual, la expresión inteligencia artificial, en su expresión amplia, no se puede ver reflejada. En nuestro sentir, el término correcto debería ser el de “simulador de inteligencia”.

Ahora bien, cualquiera que sea el término utilizado, la inteligencia artificial ha sido definida como una disciplina científica que tiene por objetivo la investigación, creación y simulación de la inteligencia humana. El término “inteligencia artificial” fue adoptado por John McCarthy, pionero americano de la inteligencia artificial.

Marvin Lee Minsky, a su vez, define la inteligencia artificial como “la construcción de programas informáticos dedicados a tareas que son, por el momento, cumplidas de manera más satisfactoria por seres humanos porque estas requieren procesos mentales de alto nivel tales como: el aprendizaje perceptivo, la organización de la memoria y el racionamiento crítico”⁵³⁵.

Constantino Ciampi sostiene que el interés de esta disciplina se reagrupa en tres sectores: El primero, el análisis y la comprensión automática del lenguaje natural.

⁵³⁵ LAROUSSE. (n.d.). Consultada en marzo 5, 2015.

http://www.larousse.fr/encyclopedie/article/Lintelligence_Artificielle/11011577

Segundo, la resolución automática del problema y la demostración del teorema. Tercero, el reconocimiento de la forma y los objetos en el espacio con aplicación a la robótica⁵³⁶.

3.5.2. La inteligencia jurídica artificial

El razonamiento jurídico no se escapa a la inteligencia artificial y la informática jurídica actual, se encuentra encaminada hacia el objetivo de construir programas informáticos, capaces de imitar mediante un proceso electrónico, la forma como se construye y se arroja el razonamiento jurídico. Tal y como en su momento, Alan Turing lo describió con su *imitation game*, la inteligencia artificial aplicada al derecho, es la imitación del raciocinio del jurista. El procesamiento, la memoria, los metadatos, las ontologías, el uso de la lógica y los algoritmos son herramientas que permiten, hoy en día, hacer hablar los datos⁵³⁷.

La aplicación jurídica de la inteligencia artificial, tiene su origen en el año 1970 cuando Buchanan y Headrick realizaron la publicación de un artículo titulado “Some speculation about artificial intelligence and legal reasoning”⁵³⁸, en el cual, por

⁵³⁶ CAMPI, C. (1982). *Intelligenza artificiale e sistema informativi giuridici*. Italia: Istituto di Teoría e Tecniche dell'Informazione Giuridica. Consultada en marzo 3, 2015. http://www.ittig.cnr.it/EditoriaServizi/AttivitaEditoriale/InformaticaEDiritto/1982_02_079-091_Ciampi.pdf

⁵³⁷ YONCK, R. (2011, Enero 17). *What is a Milestone in Artificial Intelligence?* Consultada en febrero 2, 2015. World Future Society: <http://intelligent-future.com>

⁵³⁸ BUCHANAN, B., & CKHEADRI, T. (1970, Noviembre). *Some speculation about artificial intellegence and legal reasoning*. Stanford Law Review, 23(1), p. 40. doi:10.2307/1227753
Bundesanzeiger Verlag GmbH. (n.d.). Consultada en agosto 30, 2014. <http://www.bundesanzeiger-verlag.de/ueber-uns.html>

primera vez, consideraron la posibilidad de utilizar la técnica de la inteligencia artificial como soporte al proceso de razonamiento y de decisión jurídica.

En este artículo, publicado en la Universidad de Stanford, se presentan dos razones principales por las cuales, el esfuerzo de crear programas para simular el racionamiento jurídico, podría traer consigo beneficios importantes; el primer argumento sostiene que esto podría contribuir a resolver más rápidamente, los problemas jurídicos, y el segundo, sostiene que esto podría ayudar al conocimiento de la solución del problema de las capacidades de un computador.

Las investigaciones sobre inteligencia artificial aplicadas al derecho, en la actualidad, han multiplicado los trabajos ontológicos, en filosofía del derecho. El tratamiento automático de lenguaje natural, la fusión de datos y los sistemas expertos, son todos dispositivos que intervienen unidos como un sistema de inteligencia artificial; algunos de estos ya han sido puestos en marcha, dentro de los diferentes bancos de datos jurídicos⁵³⁹.

La noción de inteligencia jurídica artificial o (*Artificial Legal Intelligence*) es una expresión creada con el fin de indicar la existencia de una ramificación de la inteligencia artificial general, que se ocupa solamente de dos sectores: el primero, el análisis y comprensión del lenguaje natural (en este campo, de inteligencia jurídica artificial, en el cual se pretende resolver el problema, hasta ahora no resuelto, del sistema informativo jurídico documental); el segundo, la resolución del problema

⁵³⁹ Como ejemplo podemos referenciar el banco de datos jurídico de Francia, el cual cuenta con tratamiento automático del lenguaje natural. Para navegar el sitio web <http://www.legifrance.gouv.fr/>

que, en el campo jurídico se traduce por la simulación del cálculo del proceso del pensamiento jurídico, del razonamiento jurídico, del proceso de decisión jurídico y de la construcción de una máquina de aplicación automática del derecho “law machine”⁵⁴⁰.

Habría que ver si en un futuro próximo o lejano, esta nueva disciplina de la inteligencia jurídica artificial culmine con la sustitución del operador jurídico, por un robot, si esta disciplina llegase a combinarse con la robótica, lo que actualmente todavía resulta una utopía.

Hasta el momento actual, la inteligencia jurídica artificial se ha ocupado solamente del estudio del lenguaje natural, que es un campo ampliamente explotado y económicamente rentable. La aplicación de fórmulas y estructuras que permiten organizar información, se considera como inteligencia artificial, toda vez que hace posible el acceso a grandes volúmenes de información, al interior de un banco de datos jurídico, con precisión de los resultados pertinentes, confrontando las palabras claves, introducidas al momento en que el usuario realiza la solicitud con las palabras de cada uno de los textos jurídicos incluidos en las bases de datos. A estos resultados es posible otorgarles un carácter semántico.

La utilización de software que permitan organizar y materializar el sistema jurídico inteligente, permite entre otras examinar la sintaxis y la base del lenguaje utilizado en el documento. No obstante, actualmente, ningún software, ninguna fórmula o

⁵⁴⁰ CAMPI, C. (1982). op. cit. p. 81.

teoría permiten encontrar, eficazmente, la información implícita, sea a nivel de léxico, a nivel de la estructura profunda o a nivel de la sintaxis lógica.

El objetivo de la inteligencia jurídica artificial, consiste en generar significados derivados, de manera que no solamente se arroje la información, en la búsqueda sino que esta información, sea interpretada por procesos lógicos de inducción o de deducción.

En este sentido, el tesoro o las estructuras arborescentes, se presentan como la base para asociar conceptos y reenviar a las citaciones necesarias del tema, asunto que, en el ámbito jurídico, es indispensable para fortalecer la argumentación y que, si se realiza de forma manual, puede presentar el riesgo de olvidar información relevante que debería ser incluida, normalmente.

De otra parte, los objetivos principales de la inteligencia jurídica artificial en el mundo del derecho, se pueden resumir en los siguientes: eficiencia en la labor, racionalización de la actividad, eficacia axiológica, autoconciencia del jurista y perfeccionamiento metodológico⁵⁴¹.

La materialización de la inteligencia artificial, en el derecho, deberá, en todos los casos, intentar resolver el problema que se plantea al representar la inteligencia humana, es decir, crear un software o una máquina capaz de percibir, comprender, predecir, modificar la información a priori y comunicarla a otro agente. Para

⁵⁴¹ SARTOR, G. (2006). *Il diritto nella societ dell'informazione*. In G. Sartor, Informatica Giuridica. Consultada en abril 30, 2015.
http://www2.cirsfid.unibo.it/didattica/upload/176_GS2005LIGTestoCap1.pdf

solucionar este problema, se hará necesaria la creación de algoritmos y combinaciones que sean susceptibles de aplicarse a los datos, mediante un sistema.

Un sistema de inteligencia artificial, aplicado al derecho, puede ser utilizado con diferentes finalidades, dentro de las cuales encontramos dos que resultan de gran utilidad, para el operador jurídico:

La primera finalidad, consiste en realizar el análisis jurídico para establecer la calificación jurídica de una situación fáctica. En este caso, nos encontramos en frente de un problema jurídico concreto al cual el sistema deberá otorgar una o varias soluciones, según el caso; es posible, igualmente, que el sistema clasifique el caso en determinada categoría jurídica; para su materialización, es idóneo el uso de sistemas expertos.

Los sistemas expertos, constituyen una de las ramas de la inteligencia artificial, en la cual los sistemas informáticos simulan el proceso de aprendizaje jurídico, memorización de la información, razonamiento jurídico y su comunicación. Con el uso y aplicación de este tipo de sistemas, es posible obtener conclusiones lógicas para realizar la toma de decisiones, atendiendo a las experiencias precedentes (tal y como lo realizaría el cerebro humano) y a la información almacenada. Pese a la importancia del tema, dentro de la inteligencia artificial, este tema no será tratado en el presente trabajo, toda vez que los sistemas expertos se ocupan de reproducir

actuaciones⁵⁴² y el tema que nos ocupa es la organización y optimización de los bancos de datos jurídicos.

La segunda finalidad, de la creación de un sistema de inteligencia artificial, es el análisis del ordenamiento jurídico, dentro del cual las fuentes formales del derecho se analizan y se ubican dentro de una estructura, en la cual sea posible interactuar, para establecer conexiones entre si y optimizar la información, para lograr que los datos contenidos en ella “hablen” dando una visión amplia y completa del tema objeto de búsqueda.

El estudio de esta finalidad, es el tema que nos ocupará en lo sucesivo del presente trabajo, la creación de un sistema que permita agrupar la información sobre la base del contenido conceptual atendiendo a su relevancia, para búsqueda conceptual de la información jurídica. Dicho en otras palabras, la creación de un banco de datos jurídicos, inteligente.

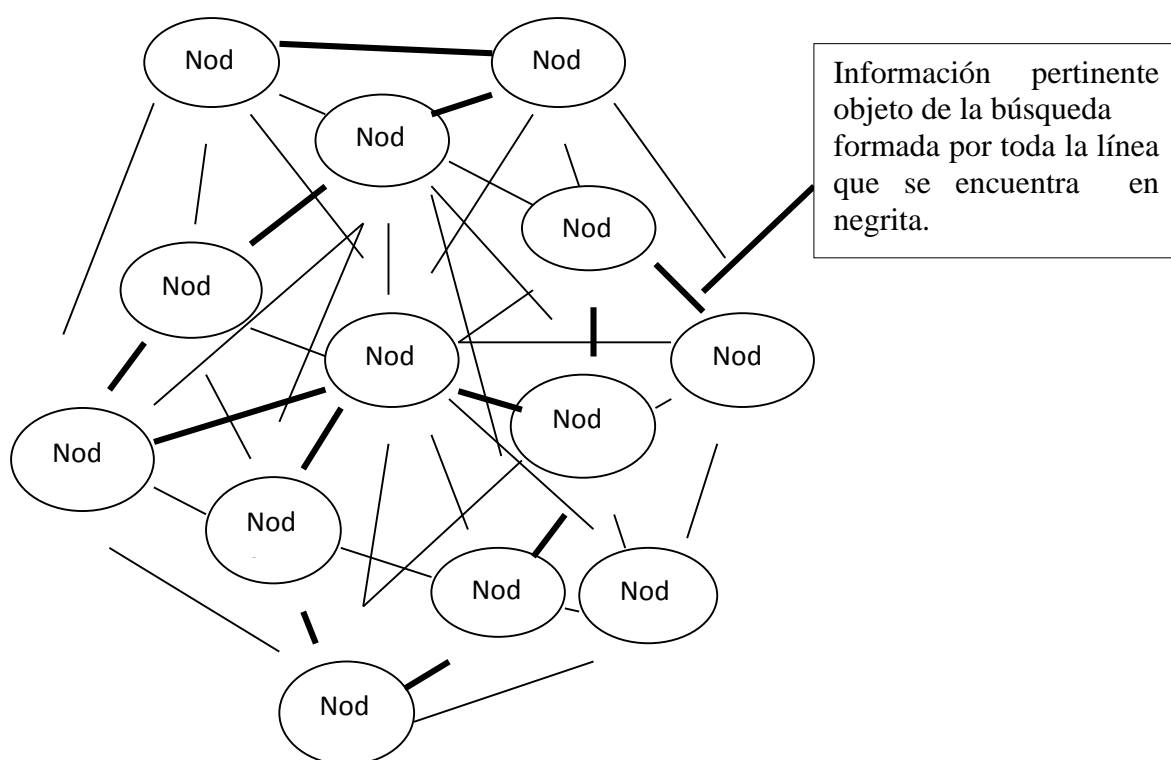
En este punto, se podría decir que para recrear la inteligencia jurídica artificialmente, se puede asimilar cada uno de los conceptos jurídicos, documentos jurídicos e información contenida en una base de datos a un nodo; estos nodos deben encontrarse interconectados, con el fin de producir lo que en el cerebro humano sería la sinapsis generada por impulsos nerviosos que viajan hasta llegar a los botones terminales de la neurona para conectarse con otras neuronas. Este proceso de “sinapsis”, en el caso que nos ocupa, se generaría gracias a la aplicación de los algoritmos y enlaces entre conceptos que una vez automatizados, se conectarán

⁵⁴² LANCHO PEDRERA, F. (2003). *Los sistemas expertos en el derecho*. Anuario de la Facultad de Derecho de la Universidad de la Rioja, XXI, p. 629-536.

creando una red de información que, al ser utilizada por el usuario, arroja la información pertinente, objeto de la búsqueda.

Lo anterior se podría representar en un gráfico de la siguiente forma:

Figura 4. Representación gráfica de inteligencia jurídica artificial



Fuente: elaboración propia con base en conceptos de inteligencia jurídica artificial

Para que el gráfico *se utilice*, es necesaria la existencia de una aplicación informática que consiste en conectar la estructura, gracias a la interacción y a su programación. La aplicación en la informática jurídica documental se ha materializado por conducto de la ontología jurídica que estudia, concretamente, la forma en que se conecta la

información jurídica por medio del estudio del lenguaje natural y de la aplicación de diferentes teorías que permiten materializar, en un software, al menos parcialmente, el raciocinio jurídico.

La idea de simular la inteligencia humana no es nueva, si bien el razonamiento jurídico no ha sido desplazado por las máquinas, los bancos de datos jurídicos constituyen una herramienta importante para ayudar a formar y robustecer dicho razonamiento.

3.6. Hacia la creación de una ontología jurídica

Dentro del desarrollo de la web semántica y de la ontología, es importante la aplicación, al derecho, de todos estos conceptos que van a incidir en el pensamiento jurídico, motivando nuevas concepciones, nuevas figuras jurídicas y permitiendo que el derecho responda adecuadamente a las necesidades de seguridad, precisión y rapidez que cada día más exige y experimenta la sociedad.

Es así como la creación de la ontología jurídica, va a permitir la construcción de un modelo apropiado para identificar la información en forma adecuada y mantenerla permanentemente actualizada.

3.6.1. El ordenamiento Jurídico como un sistema complejo

El ordenamiento jurídico es un sistema complejo que se encuentra compuesto por varias partes interconectadas entre sí, gracias a la interpretación que realiza el

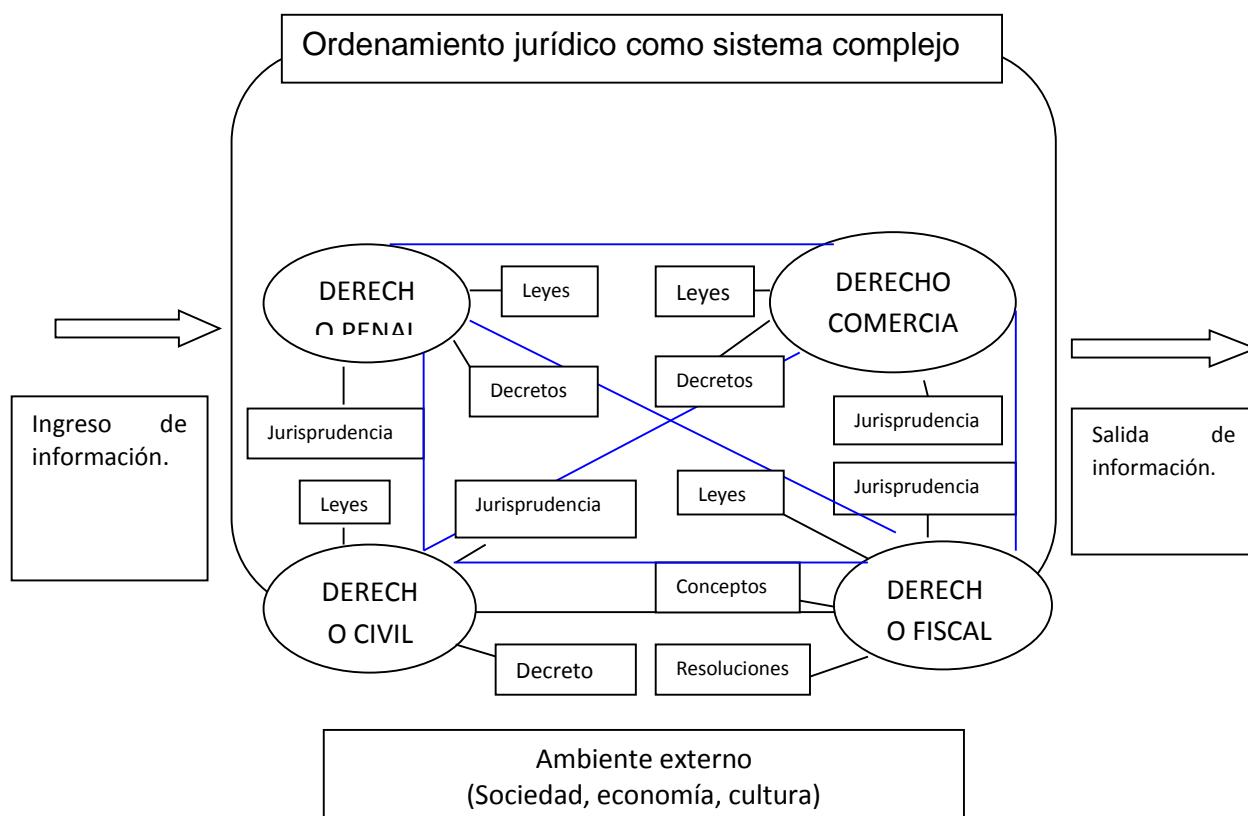
operador jurídico, estas interconexiones dan nacimiento a información adicional que nace una vez se analiza el conjunto de datos jurídicos y que a priori no son visibles antes por el observador. Los documentos jurídicos interactúan entre sí y crean nuevas informaciones jurídicas valiosas, que no se pueden interpretar ni explicar a partir de un documento jurídico aislado sino en su conjunto.

La interpretación completa y el análisis de una jurisprudencia que decide sobre la constitucionalidad de una ley, no pueden realizarse de manera aislada. Para poder entender este documento jurídico, es necesario interpretarlo en conjunto con la constitución, la Ley objeto de estudio e incluso es necesario recurrir a otras áreas del derecho para comprender el alcance y la totalidad de sus implicaciones.

En otras palabras, el ordenamiento jurídico como sistema complejo que es, contiene un sin número de documentos jurídicos (leyes, códigos, decretos, jurisprudencia, resoluciones, conceptos, tratados, etc.), que varían de un área del derecho a la otra, pero que en muchas ocasiones son interdependientes y regulan otras áreas del derecho. Un ejemplo claro podría ser una ley que reforma el sistema tributario y que incluye una reforma en materia de derecho de sucesiones. El ordenamiento jurídico como sistema complejo, posee mayor información de la que puede obtenerse si se analiza un documento jurídico independiente; para conocer y poder interpretar verdaderamente este sistema complejo, será necesario no solamente conocer el funcionamiento de una ley sino del ordenamiento jurídico completo, relacionando todos los documentos entre sí.

Los documentos jurídicos no pueden ser interpretados aisladamente, porque el derecho es transversal, a pesar de analizar un documento especializado en un área del derecho, será indispensable conocer el conjunto del ordenamiento jurídico para que este documento aislado se incorpore y se articule con todo el engranaje jurídico, haciendo funcionar el sistema en su conjunto. Este sistema complejo puede ser representado en el siguiente gráfico:

Figura 5. Ordenamiento jurídico como sistema complejo



Fuente: elaboración propia con base en conceptos del ordenamiento jurídico como sistema complejo

Los sistemas complejos se caracterizan porque su comportamiento es imprevisible, verbigracia, un cambio de jurisprudencia o la derogatoria de una ley, puede ser la

respuesta a un fenómeno social, económico, político, cultural, etc. Esta imprevisibilidad hace necesario que se estudie e incluya dentro de una ontología jurídica, el análisis del ordenamiento jurídico como sistema complejo; complejidad no en el entendido de complicación o de dificultad en la comprensión. De igual forma, el ordenamiento jurídico como sistema complejo es interdependiente, conectado, diverso y adaptativo.

Dentro del ordenamiento jurídico existe igualmente una influencia mutua entre sus elementos⁵⁴³, una ley que regula la protección al consumidor puede influenciar la regulación en materia de comercio electrónico, toda vez que el cambio en uno de los documentos que forman el sistema complejo, repercute inevitablemente en el resto. Para comprender en su totalidad este sistema, es necesario conocer las interacciones o relaciones entre todos y cada uno de los documentos jurídicos, así como las evoluciones y cambios.

Cabe afirmar en igual sentido, que el ordenamiento jurídico como sistema complejo es abierto, por cuanto el mismo derecho es producto de la interacción social y que es justamente el entorno lo que define las diferentes tendencias, porque el derecho reconoce las diferentes situaciones sociales y las regula desde sus diferentes ámbitos. Estos cambios sociales, económicos o culturales crean un dinamismo constante en el interior del sistema complejo objeto de estudio, ya que éste se encuentra sumiso en un entorno variable que influye de manera considerable y genera imprevisibilidad en sus tendencias jurídicas.

⁵⁴³ MORIELLO, S. A. (2013, Septiembre 13). *Investigación: Sistemas complejos*. Consultada en marzo 2, 2015. <http://www.cs.us.es/~fsancho/?p=sistemas-complejos-2>.

El ordenamiento jurídico cumple con los tres requisitos para ser considerado un sistema complejo, a saber:

*“En primer lugar, **está compuesto por una gran cantidad de elementos relativamente idénticos**. Por ejemplo, las células en un organismo, o las personas en una sociedad.*

*En segundo lugar, **la interacción entre sus elementos es local** y origina un comportamiento emergente que no puede explicarse a partir de dichos elementos tomados aisladamente. Un desierto puede contener billones de granos de arena, pero sus interacciones son excesivamente simples comparadas con las que se verifican en las abejas de un enjambre.*

*Por último, es **muy difícil predecir su evolución dinámica futura**; o sea, es prácticamente imposible vaticinar lo que ocurrirá más allá de un cierto horizonte temporal. (...)*

Cualquier variación mínima entre sus elementos componentes puede modificar, de forma imprevisible, las interrelaciones y, por lo tanto, el comportamiento de todo el sistema. Así, la evolución de esta clase de sistemas se caracteriza por la fluctuación, situación en la que el orden y el desorden se alternan constantemente”⁵⁴⁴. Negrillas dentro del texto original.

⁵⁴⁴ Ibídem.

La complejidad de los documentos jurídicos y la necesidad de organizar el conocimiento, generan como consecuencia la necesidad de utilizar herramientas, capaces de cumplir con el objetivo de articular y organizar la información de forma coherente y de asimilar el derecho a un sistema complejo. En otras palabras, la disciplina jurídica exige herramientas capaces de tratar el derecho como sistema complejo, porque un corpus jurídico debe generar su propia dinámica. En consecuencia, todo proceso de creación de ontologías jurídicas, deberá incorporar la complejidad de las relaciones que concatenan no solamente los términos entre sí sino la ontología y el corpus⁵⁴⁵.

3.6.2. Definición de la ontología jurídica

Se encuentra en la actualidad, no solo en el ámbito jurídico, bases de datos creadas con un objetivo que va mucho más allá de la simple recopilación de información⁵⁴⁶. Estas bases de datos además de organizar el documento, según la temática, lo sitúa en un contexto determinado, de manera que el usuario puede obtener respuestas más completas, analizadas desde varias perspectivas gracias a la creación de enlaces entre sus contenidos.

En las bases de datos jurídicas, existe la más de las veces un análisis de los documentos para determinar la tesis principal del documento e identificar los temas

⁵⁴⁵ MAZZEGA, P., BOUCIER, D., BOURGINE, P., & NADAH, N. (n.d.). *Ontologie juridique et systemes complexes*. Consultada en febrero 4, 2015.

<http://www.gemas.fr/dphan/rochebrune09/papiers/BoucierDanielet.pdf>

⁵⁴⁶ La Universidad Externado de Colombia, en su Departamento de Derecho Informático, cuenta con bases de datos jurídicas de diferentes áreas del derecho, que incluyen el análisis de los documentos, con indicación de las tesis principales y sus problemas jurídicos y que prontamente estarán en Internet dentro de la Web 3.0.

relevantes a los cuales se hace mención, igualmente, en el caso de la jurisprudencia, es posible encontrar en una base de datos el planteamiento de uno o varios problemas jurídicos, según el caso en concreto. Como resulta lógico, este proceso requiere de personal especializado en la materia, desde la tarea de recopilación hasta el análisis. A pesar de contar con el apoyo de herramientas informáticas, reconoce los limitantes de las mismas, motivo por el cual, no deja su tarea asignada únicamente fórmulas matemáticas, toda vez que la naturaleza de las bases, exige un análisis especializado.

Como ya se ha mencionado en la primera y segunda parte, estas bases de datos cuentan con *tesauros jurídicos* especializados y constituyen la base para crear ontologías aplicadas al derecho, para evolucionar hacia la llamada *Web semántica o Web 3.0*. Lo que antes para nosotros era una utopía, hoy en día se hace posible gracias a la evolución de la tecnología. La denominada Web Semántica, va a permitir que los abogados localicen la información desde diferentes perspectivas, lo cual es un gran apoyo en el quehacer jurídico diario. Los bancos de datos jurídicos que se hallen inmersos en la Web 3.0, van a permitir encontrar la representación de un conjunto de conceptos y su relación entre los mismos.

Esta es la era que aguarda por nosotros en un futuro inmediato, denominada como la Web 3.0 o Web semántica y que surge como respuesta a las grandes cantidades de información, que ha dejado el uso de la Web 2.0 o Web Social. En términos prácticos, Web Semántica es la Web del significado, del sentido y de la interpretación del significado que se le puede otorgar en el caso que nos concierne, a un documento jurídico. Así pues, esta es la era de la red de datos o datos enlazados.

Ahora bien, dentro de la *Web semántica*, podemos encontrar la creación de lo que se ha denominado *ontologías jurídicas*. El término ontología, hace referencia a la comprensión común de un dominio de interés, que puede ser utilizado como un marco unificador destinado a resolver los problemas de comprensión compartida de información. Este término a su vez, ha sido traído de la filosofía, donde es entendido como una teoría de la naturaleza de la existencia. Al adaptarlo al campo de Web, se entiende como un documento u archivo que define formalmente la relación entre términos⁵⁴⁷.

Para la materialización de la Web Semántica, es necesario recurrir a la Ontología, la cual se puede definir como: “una especificación explícita y formal sobre una conceptualización compartida. Las ontologías permiten definir de manera compartida y consensuada conceptos sobre un área específica del conocimiento y establecer las relaciones entre dichos conceptos. Esto permite a humanos y máquinas comunicarse de manera eficiente al proveer una definición común de un dominio”⁵⁴⁸.

La ontología viene a ser una conceptualización o estandarización dentro de un campo de conocimiento.

Una ontología “es una jerarquía de conceptos con atributos y relaciones que define una terminología consensuada para definir redes semánticas de unidades de información interrelacionadas. Proporciona un vocabulario de clases y relaciones para describir un dominio, enfatizado en la comparación del conocimiento y el

⁵⁴⁷ BERNERS-LEE, T., HENDLER, J., & LASILLA, O. (2001). op. cit.

⁵⁴⁸ GUZMAN LUNA, J., LÓPEZ GARCÍA, A. N., & TORRES PARDO, D. (2006). *Desarrollo de una ontología en el contexto de la web semántica a partir de un tesoro documental tradicional*. Revista Interamericana de Bibliotecología, 29(2), p. 79.

consenso en su representación”⁵⁴⁹.

Una ontología necesariamente implica o representa algún tipo de visión del mundo amplia con respecto a un determinado tema. La visión del mundo a menudo se concibe como un conjunto de conceptos, definiciones e interrelaciones, lo que se conoce como la conceptualización del conocimiento.

“Debe resaltarse que el término ontología no se utiliza en la misma acepción que en la disciplina de lógica ni tampoco en el sentido tradicionalmente utilizado en Filosofía del Derecho, el ‘ser’, sino con otro sentido. Señalan Russell y Norvig que la ontología general es una representación que organiza todas las cosas del mundo mediante una jerarquía de categorías. Ciertamente que dicha representación es en ocasiones difícil debido a la existencia de excepciones o de eventos que mutan los objetos. Como ejemplo de excepción podemos citar cualquier categoría que pudiera estar subordinada a varios objetos simultáneamente (por ejemplo, el ánimo de lucro tiene diferentes sentidos según sea un concepto civil o penal, por lo que la ontología deberá prever los términos polisémicos). Como ejemplo de la dificultad que supone la existencia de eventos, un imputado puede mutar en un acusado y posteriormente en un condenado⁵⁵⁰”.

⁵⁴⁹ ESPINEL, Á., GUEVARA BOLAÑOS, J. C., FLÓREZ FERNÁNDEZ, H., PÉREZ CASTILLO, J. N., & PINZÓN NUÑEZ, S. A. (2006). La web semántica y sus posibles aplicaciones en el ámbito universitario. *Revista Tecnura*(19), p. 121-131.

⁵⁵⁰ DE LA CUEVA, J. (2011, Junio 11). *Proyecto Kelsen*. <http://derecho-internet.org/proyectos/documentacion/kelsen/teoria/informacion/ontologia>

En el ámbito jurídico, la ontología tiene por finalidad obtener un concepto jurídico para posteriormente ser utilizado a manera de reflexión, lo anterior se obtiene mediante el análisis y la correlación de cada uno de los componentes del derecho, más concretamente, del ordenamiento jurídico. De igual forma, las ontologías jurídicas tienen como finalidad mejorar los cuestionamientos dentro de una base de datos jurídica.

Aplicado al derecho, una ontología jurídica pretende estudiar la forma del conocimiento jurídico con el objetivo de representarlo. De igual manera, el término ontología hace referencia al estudio y la representación de conocimiento, que pueden ser manejados en un software. En el ámbito jurídico, las aplicaciones pueden ser tanto teóricas como prácticas.

Gracias a su implementación es posible analizar las causas que originan el racionamiento jurídico y generar conocimiento. Mediante la creación, la aplicación y el uso de una ontología, es posible simular el raciocinio jurídico y examinar las causas que enlazan el conocimiento. Lo anterior, se logra con la creación y representación de las relaciones, por ejemplo, entre los aspectos generales de un documento jurídico y lo particular del mismo.

Para asegurarnos de los que los diferentes usuarios tengan una comprensión común de los términos, uno requiere ontologías bajo las cuales esos términos estén descritos y que permita por ende establecer una terminología común entre los usuarios. De allí que “la idea es que la Web Semántica esté formada (al menos en parte) por una red de nodos tipificados e interconectados mediante clases y relaciones definidas por

una ontología compartida por sus distintos autores”⁵⁵¹.

La ontología aplicada a la Web, debe ofrecer una buena comprensión de lo que está representando. Es decir, debe incluir relaciones de términos y conceptos, sus definiciones o significados y sus relaciones con cada uno de ellos. Esto permite una uniformidad y homogenización de los modelos aplicados en la Web⁵⁵².

Las ontologías pueden tomar variedad de formas, pero necesariamente deberán incluir un vocabulario de términos y alguna especificación de su significado o definición. Por esta razón, la base para la creación de una *ontología jurídica* son los tesauros jurídicos, que deberán incluir dentro de los mismos, un conjunto de descriptores definidos rigurosamente con la semántica jurídica formal, que atienda al máximo grado de formalidad del vocabulario⁵⁵³. Lo expuesto con anterioridad, se elabora con el fin de centralizar los resultados de la búsqueda en atención a los diferentes términos que pueden llegar a hacer referencia al mismo documento objeto de la búsqueda. La ontología ofrecerá así, la semántica a los bancos de datos jurídicos, que deberán desarrollarse de manera cuidadosa, exhaustiva y apoyada por un grupo de profesionales experimentados en cada área del derecho.

⁵⁵¹ CASTELLS, P. op. cit.

⁵⁵² FLOREZ FERNÁNDEZ, H. (2006, Octubre 4). *Web Semántica aplicada al registro académico Institucional*. Consultada en mayo 23, 2015.

http://www.konradlorenz.edu.co/images/stories/suma_digital_sistemas/2009_01/Informe_Investigacion_semantica.pdf

⁵⁵³ “El grado de formalidad con el cual se crea un vocabulario y el significado especificado es muy variable. Cuatro son las formas de expresión de vocabulario más frecuentes a saber: -Muy informal: Es expresado libremente en el lenguaje natural. -Semi-informal: Se expresa en una forma restringida y estructurada del lenguaje natural. -Semi-formal: Reexpresadas en un lenguaje artificial formalmente definido, por ejemplo, la versión de la ontología Ontolingua enterprise. - Rigurosamente formal: Términos meticulosamente definidos con la semántica formal, teoremas y demostraciones de propiedades tales como la solidez y la integridad. Por ejemplo: El modelo TOVE de la Universidad de Toronto Canadá”.

Las ontologías jurídicas se desarrollaron inicialmente dentro del programa de la inteligencia artificial, como una tecnología de representación del contenido legal — Frame Based Ontology (FBO), LRI Core, Legal Knowledge Interchange Format (LKIF). Sin embargo, en el actual panorama de su incorporación a Internet, ha crecido el interés en su aplicación a dos aspectos más específicos, íntimamente relacionados: la recuperación de la información basada en vocabularios controlados y la organización de sitios web mediante taxonomías. En este sentido, muchos de los proyectos de los últimos años se han centrado en la expresión en OWL de tesauros o taxonomías legales de alcance general o especializado⁵⁵⁴.

El tesauro, es así un sistema de organización del conocimiento, cuyos términos de indexación se estructuran en un sistema de relaciones, que consta de un vocabulario de indexación conformado por descriptores e identificadores. Ambos, son términos aceptados en la indexación para describir el contenido de los documentos en los catálogos de la biblioteca, bibliografías u otros documentos secundarios. “Su función es remitir desde un término no autorizado al término equivalencia, cuyos descriptores, identificadores y no descriptores forman el vocabulario de entrada del tesauro. En un tesauro, se hacen explícitos tres tipos de relaciones:

- Relaciones de equivalencia.
- Relaciones de jerarquía.

⁵⁵⁴ SCHWEGHOFFER, E., & LIEBWALD, D. (2008, Junio). *Advanced lexical ontologies and hybrid knowledge bases systems: first steps to a dynamic legal electronic commentary*. Artificial Intelligence and Law, 15(2), p. 103-115.

- Relaciones de asociación”⁵⁵⁵.

En este orden de ideas, las *ontologías* pretenden evitar dificultades en la identificación del objeto de la búsqueda y variación de los resultados con relación a la información solicitada, en el caso de emprender búsquedas con diferentes palabras clave. La manera de abordar y solucionar estos problemas, es reducir o eliminar la confusión conceptual y terminológica, con el fin de llegar a un entendimiento compartido.

Tal comprensión puede funcionar como marco unificador para los diferentes puntos de vista y sirve como base para generar comunicación y proporcionar la información completa a los abogados, teniendo presentes las diferentes necesidades y puntos de vista derivados de sus diferentes contextos.

Finalmente, es importante mencionar que en un futuro próximo nos vamos a dirigir hacia bancos de datos y buscadores cada vez más semánticos, con capacidad de entender preguntas y responder con soluciones correctas a los diferentes planteamientos, en el caso que nos interesa de la informática jurídica, podremos encontrar una correspondencia a situaciones que serán descritas y solucionadas por los cada vez más evolucionados sistemas de información.

⁵⁵⁵ GARCÍA JIMÉNEZ, A. (2004). *Instrumentos de representación del conocimiento: tesauros versus ontologías*. Anales de la Documentación, 79-95. <http://revistas.um.es/analesdoc/article/view/1691>

3.6.3. La creación de una ontología jurídica

Existen nuevas formas de construir ontologías jurídicas a partir del lenguaje natural. Los modos de construcción actuales de una ontología, obligan a crear una estructura particular del conocimiento y un número clausus de relaciones.

Es importante tener presente que el ordenamiento jurídico es un sistema complejo y dinámico, lo que genera limitaciones a la hora de crear una ontología jurídica. Estas limitaciones restringen la evolución, sin embargo, es posible encontrar nuevas herramientas para conservar, medir y visualizar la complejidad de las relaciones entre el ordenamiento jurídico y la ontología. Las nuevas perspectivas realizan una alianza entre la web semántica (ya explicada con anterioridad) y algunas propiedades de los sistemas complejos, dentro de la cual se podrá crear una interacción entre el ordenamiento jurídico y el operador jurídico.

El uso de las ontologías se inscribe en la imperiosa necesidad de organizar y recuperar todo el conjunto de documentos jurídicos, que componen el sistema complejo del ordenamiento jurídico de forma automatizada. Ahora bien, la particularidad de las ontologías jurídicas reside en que para definir las relaciones entre el conjunto de documentos jurídicos, es indispensable contar con el conocimiento y el criterio jurídico.

El ordenamiento jurídico y en particular cualquier base de datos jurídica actual, requiere no solamente relacionar los términos entre sí, sino también integrar los

documentos, uniendo la jurisprudencia con la legislación y la doctrina para generar una visión completa, en cada asunto concreto.

3.7. Ordenamiento jurídico: sistema complejo con uso de lenguaje natural

Nos corresponde en esta parte analizar las características que comparten el ordenamiento jurídico con los sistemas complejos y más concretamente, qué soluciones podrían ser puestas en marcha para lograr enlazar la ontología con su inmersión en un cuerpo creado en lenguaje natural. Esas herramientas pueden ser utilizadas de forma simultánea a la construcción misma de la ontología.

En el lenguaje natural aplicado al derecho, se hace necesario tener presente las evoluciones jurídicas y los cambios de interpretación. “Los programas y los filósofos del derecho han desarrollado sistemas conceptuales muy anteriores a la elaboración de las ontologías lingüísticas, las investigaciones sobre la inteligencia artificial aplicada al derecho han multiplicado los trabajos ontológicos en la teoría del derecho”⁵⁵⁶.

Ahora bien, para realizar el trabajo lingüístico, es indispensable realizar un trabajo multidisciplinario y conjunto entre un jurista especializado y un ingeniero informático programador, toda vez que es indispensable representar los conocimientos de manera automatizada. Estos conocimientos son compartidos, en gran medida, gracias al uso del lenguaje natural del documento jurídico, a saber: la totalidad del texto, el proceso hermenéutico, el uso de palabras explícitas, etc. Lo

⁵⁵⁶ MAZZEGA, P., BOUCIER, D., BOURGINE, P., & NADAH, N. op. cit.

anterior, constituye una de las bases fundamentales para poder desarrollar una ontología, ya que es indispensable representar la complejidad del sistema en un software, porque el objetivo fundamental, es lograr que los documentos contenidos en el banco de datos, hablen.

Ahora bien, existen tres aspectos importantes que hay que tener en cuenta a la hora de elaborar una ontología jurídica:

El primero consiste en examinar las fuentes y el texto de cada documento, una vez realizada la automatización a partir del texto.

En segundo lugar, es recomendable verificar si existe o no una estructura arbórea dentro del documento jurídico. Tratándose de documentos jurídicos codificados, se encuentra una estructura arbórea presente, que de no ser así, es recomendable elaborarla, toda vez que el tesaurus o la estructura arbórea muestra la taxonomía del documento jurídico, lo cual, es fundamental para comprender el contexto de cada individualidad. Esto nos permite además, tener una mayor cercanía con el raciocinio jurídico y crear relaciones entre las diferentes estructuras arbóreas.

Finalmente, otro aspecto fundamental, consiste en implementar técnicas para establecer conceptos, relaciones entre documentos y dentro del documento mismo, teniendo en cuenta su jerarquía. Este análisis de sintáctica del texto busca establecer el significado de cada uno de sus componentes y elegir dentro de ellos, cuales deberán formar parte de los conceptos a incluir en la ontología. Una vez realizado el análisis de la terminología, es posible construir la ontología a partir de los extractores

de término⁵⁵⁷. “Algunas herramientas establecen enlaces entre los recursos termino-ontológicos y el corpus. Desafortunadamente la gran mayoría de los editores de ontologías no suministran los medios para conservar los enlaces desarrollados durante el proceso de desarrollo”⁵⁵⁸.

Este proceso de examen terminológico, resulta fundamental para la creación de una ontología jurídica, que realmente contenga la semántica y permita acercarse a la forma de un raciocinio jurídico. Sobre este punto, es posible recurrir a las reglas jurídicas que ya han sido formalizadas y definidas por reconocidos expertos o por los organismos competentes.

El éxito de una ontología jurídica, radica en la posibilidad de automatizar estas tres diferentes formas de tratar la información jurídica que se incluye en un banco de datos. El proceso más complejo, consistirá en representar en un software los enlaces desarrollados por el experto jurídico, lo que se encuentra a cargo del experto técnico en informática.

Para construir una ontología mediante el tratamiento automático del lenguaje natural, es pertinente hacer referencia a la autora francesa Danièle Bourcier, quien elaboró la base de datos jurídica de *legifrance*, que contiene la totalidad del ordenamiento jurídico francés, en línea y con acceso libre, base que ya ha sido abordada y analizada en la segunda parte de esta tesis.

⁵⁵⁷ Ibidem.

⁵⁵⁸ Ibidem.

Ella sostiene que existe un método fundado sobre el concepto de ontología diferencial, cuyo objetivo radica en buscar el nivel pertinente para crear una ontología genérica entre los formatos existentes y los sistemas jurídicos. Para lograr este objetivo, sostiene la autora, “es necesario trabajar a partir de los estándares existentes utilizando los *use cases*⁵⁵⁹ y las reglas jurídicas ya formalizadas por los expertos”⁵⁶⁰.

Estos estándares existentes son los denominados RELs, que se elaboran a partir de los diccionarios de datos (*Right Data Dictionaries RDD*)⁵⁶¹. Los diccionarios, no son fundados ni sobre las ontologías ni sobre el derecho. Por lo cual, es necesario definir las nociones precisas, para crear la ontología, otorgando una definición para cada caso en lenguaje natural.

Ahora bien, debido a la peculiaridad de las fuentes jurídicas, en la gran mayoría de oportunidades, resultará imposible construir una ontología mediante el uso de un sistema automático o semiautomático, toda vez que los documentos jurídicos

⁵⁵⁹ En software e ingeniería de sistemas, un *use case* consiste en una lista de pasos que define las interacciones típicas entre un rol (conocido en el Lenguaje Unificado de Modelado UML, por sus siglas en inglés Unified Modeling Language como un “actor”, el actor a su vez especifica el rol jugado por un usuario o por un sistema cuando estos dos interactúan entre sí sobre un tema) y un sistema, con el fin de lograr un objetivo. El actor puede ser un humano, un sistema externo o el tiempo.

⁵⁶⁰ MAZZEGA, P., BOUCIER, D., BOURGINE, P., & NADAH, N. op. cit.

⁵⁶¹ “Un diccionario de datos es un conjunto de metadatos que contiene características lógicas y puntuales de los datos que se van a utilizar en el sistema que se programa incluyendo nombre, descripción, contenido y organización. Es un catálogo, un depósito, de los elementos en un sistema. Como su nombre lo sugiere, estos elementos se centran alrededor de los datos y la forma en que están estructurados para satisfacer los requerimientos de los usuarios y las necesidades de la organización. En un diccionario de datos se encuentra la lista de todos los elementos que forman parte del flujo de datos en todo el sistema. Los elementos más importantes son flujos de datos, almacenes de datos y procesos. El diccionario guarda los detalles y descripciones de todos estos elementos”. http://es.wikipedia.org/wiki/Diccionario_de_datos

cuentan con diferentes particularidades, de manera que en este caso será posible construir manualmente la ontología.

Una vez construida la ontología, es posible editarla utilizando un *Differential Ontology Editor (DOE)*, que es un simple editor de ontología, que permite al usuario edificar las mismas, de acuerdo con la metodología propuesta por *Bruno Bachimont*⁵⁶².

El proceso se encuentra dividido en tres pasos: en el primer paso, el usuario está llamado a construir taxonomías de conceptos y relaciones, justificando expresamente la posición de cada noción en la jerarquía. Para cada noción, el usuario construye una definición. Por lo tanto, el usuario debe especificar porque una noción es similar pero más específica que otras, y porque esta noción es similar pero diferente de sus asociadas. Igualmente, puede adicionar sinónimos y definiciones de enciclopedias (que para el caso de un banco de datos jurídico, serían las definiciones contenidas en las fuentes formales del derecho y en las auxiliares, como la doctrina), para todas las nociones⁵⁶³.

En un segundo paso, las taxonomías fueron consideradas desde una extensión del punto de vista semántico. El usuario puede aumentar esto con nuevas definiciones o añadir restricciones en los dominios de las relaciones. Finalmente, en un tercer paso, la ontología puede ser traducida en un lenguaje de representación del conocimiento. Esto permite utilizarlo en un sistema basado en ontologías apropiado o importarlo en

⁵⁶² DOE - *The Differential Ontology Editor*. (2008, Junio 24). Consultada en marzo 3, 2015.
<http://www.eurecom.fr/~troncy/DOE/>

⁵⁶³ *Ibíd.*

otra herramienta de creación de la ontología como: RDFS, OWL, DAML+OIL, OIL, CGXML⁵⁶⁴. Estas herramientas son empleadas por los ingenieros de sistemas encargados de plasmar la ontología.

Es importante precisar que *Differential Ontology Editor* (DOE), no es un entorno de desarrollo de ontologías completo, de manera que DOE, no apoya activamente todas las actividades que se encuentran involucradas en la construcción de la ontología. Sin embargo, DOE actúa como un complemento de otros editores, ofreciendo técnicas de inspiración lingüística, (que se unen a las definiciones de léxico de los conceptos y las relaciones utilizadas) y justifica sus jerarquías desde un punto de vista teórico, comprensible desde el punto de vista humano⁵⁶⁵.

Ahora bien, para construir la ontología a partir del DOE (*Differential Ontology Editor*), es indispensable contar con una guía de estructuración de conceptos. ARCHONTE (ARCHitecture for ONTological Elaborating) se presenta así como una ontología “lingüística”, construida sobre la base de una semántica general⁵⁶⁶.

Aquí la metodología consiste en utilizar cuatro principios diferenciales, otorgando a cada concepto sus similitudes y disimilitudes. Verbigracia: El concepto de “ascendientes” y el concepto de “descendientes”. Aquí habrá que identificar las similitudes con el concepto de “ascendientes”, las similitudes con el concepto de “descendientes” y la diferencia con el concepto de “descendientes”. Esta

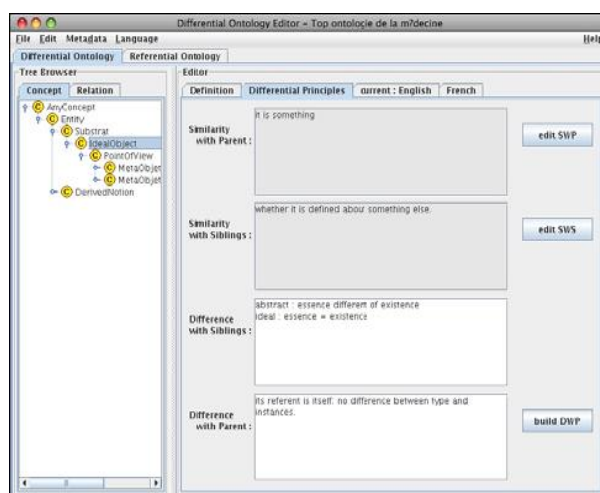
⁵⁶⁴ Algunas de estas herramientas ya han sido explicadas en el acápite concerniente a las herramientas de la web semántica

⁵⁶⁵ DOE - *The Differential Ontology Editor*. (2008, Junio 24). op. cit.

⁵⁶⁶ BACHIMONT, B. (2004). *Art et sciences du numérique: Ingénierie des connaissances et critique de la raison computationnelle*. Université de Technologie de Compiègne

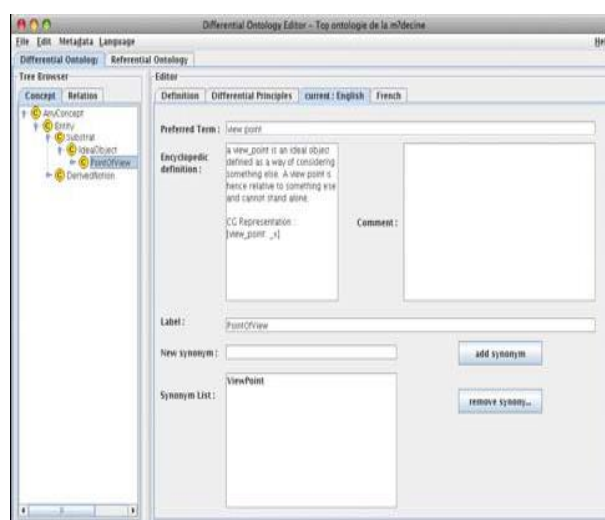
información es otorgada en lenguaje natural; la ontología resultante es denominada “ontología diferencial” y se sitúa en un nivel lingüístico⁵⁶⁷. Lo anterior, se puede ver representado en los siguientes gráficos:

Figura 6. Principios Diferenciales en DOE. (Differential Ontology Editor)



Fuente : gráfico tomado de: Ontologie juridique et systèmes complexes⁵⁶⁸.

Figura 7. Informaciones Lingüísticas en DOE. (Differential Ontology Editor)



Fuente : gráfico tomado de: Ontologie juridique et systèmes complexes⁵⁶⁹.

⁵⁶⁷ MAZZEGA, P., BOUCIER, D., BOURGINE, P., & NADAH, N. op. cit.

⁵⁶⁸ Ibidem.

3.7.1. Tratamiento del ordenamiento jurídico como sistema complejo

En un capítulo anterior explicamos cómo el ordenamiento jurídico puede ser considerado un sistema complejo, retomando esta idea, es necesario ahora materializarla o ponerla en práctica desde el punto de vista de la creación de una ontología.

La complejidad de un sistema se puede definir por cuatro grandes criterios:

“a. El gran número de componentes (artículos,...) que interactúan entre ellos en diferentes escalas;

b. el conocimiento parcial de los vínculos (dinámicos) entre los componentes que el experto humano puede tener;

c. la dificultad de prever la evolución de tal sistema y/o de los outputs;

d. la dependencia de las propiedades del sistema cuando se observan en los diferentes niveles de su organización”⁵⁷⁰.

El ordenamiento jurídico como sistema complejo, presenta dificultades a la hora de ser representado dentro de una ontología jurídica, dada su complejidad. El razonamiento de un jurista se encuentra atado a variables que dificultan la creación de un sistema capaz de representar exactamente el raciocinio jurídico; éstas variables

⁵⁶⁹ Ibidem.

⁵⁷⁰ Ibidem.

son entre otras: la inestabilidad del ordenamiento, que día a día introduce nuevas reformas que cambian la forma en que se estructura el pensamiento jurídico sobre un asunto concreto, verbigracia: un cambio de jurisprudencia, la derogación de una ley, el cambio de un término jurídico como el que ocurrió en el año 2014 con la desaparición del término “buen padre de familia” del código civil francés”⁵⁷¹.

Otra variable que dificulta la representación exacta del ordenamiento jurídico como sistema complejo, es el carácter de imprevisible. Un jurista puede prever ciertas reformas en el ordenamiento que espera se lleven a cabo, no obstante, incluso el operador jurídico humano, se encuentra imposibilitado para prepararse y conocer de manera cierta, cuál será la evolución del derecho. Esta es la razón por la cual dentro de la ontología, se encontrará siempre presente el elemento de la imprevisibilidad, como un componente defectuoso que no asegura la fiabilidad a largo plazo de la representación del sistema.

El ordenamiento jurídico se encuentra atado a la sociedad y a los cambios que se puedan producir en ella, especialmente en materia de derecho de familia, verbigracia: en temas como el aborto, el matrimonio entre parejas del mismo sexo, la eutanasia, las técnicas de fecundación humana asistida, el alquiler de vientre, etc. de igual forma, en asuntos ligados a la economía fluctuante que se encuentra supeditada a factores externos del mercado, que son controlados por las grandes potencias mundiales, estos factores económicos pueden obligar a un país a cambiar su política monetaria, a crear reformas en materia fiscal, aduanera o cambiaria.

⁵⁷¹ Ley para la igualdad real entre hombres y mujeres, publicada el 4 de agosto de 2014. El 21 de enero de 2014, la Asamblea Nacional de Francia adoptó una enmienda en la cual se suprimió la expresión « buen padre de familia » del ordenamiento jurídico francés y particularmente del Code Civil y remplazándola por el término “razonable” o “razonablemente”.

Una última variable que podríamos citar, es la actualización limitada en el tiempo, de la ontología, que es producto de las tendencias cambiantes del derecho. La ontología se elabora a priori, con las fuentes documentales presentes al momento de su creación, de manera que se desconoce cuál va a ser su evolución. Esto presenta una dificultad a la hora de reconfigurar los textos que se encuentran ya inmersos en la ontología, cuando se realiza la actualización pues, como ya se ha explicado el conocimiento deberá ir conectado y el cambio de un texto jurídico puede acarrear reformas, en otros documentos.

Ahora bien, pese a las limitaciones citadas anteriormente, es posible realizar un tratamiento del ordenamiento jurídico como un sistema complejo; para tales efectos la idea de interconexión juega un papel fundamental, si se tiene en cuenta que el conocimiento jurídico se construye en gran parte entrelazando cada uno de sus componentes de manera lógica, esto genera conocimiento.

Para materializar el tratamiento jurídico del sistema complejo que es el ordenamiento jurídico, será necesario crear interconexiones entre los documentos jurídicos que componen el banco de datos jurídicos y los contenidos cognitivos⁵⁷². Para resolver el problema de la inestabilidad jurídica, se puede proponer la definición de términos genéricos, no obstante será necesaria la actualización de la ontología cada vez que algún término jurídico varíe.

El uso de estructuras arbóreas generales, puede contribuir a resolver los problemas de inestabilidad jurídica, una buena guía de la construcción del tesoro se encuentra

⁵⁷² BOUCIER, D., & MAZZEGA, P. op. cit.

en la codificación, los cambios se introducirán dentro de la estructura arbórea y en algunos casos será necesario introducir nuevos descriptores, sin embargo como se realizaba con las bases de datos jurídicas en el pasado, la creación de nuevos descriptores dentro de un tesoro, no sucede de forma frecuente si su estructura goza de coherencia y se encuentra bien elaborada, por lo cual el cambio podrá incluso subsumirse dentro de la estructura arbórea prediseñada, salvo en los casos en que realmente exista una nueva figura jurídica o una creación que así lo amerite.

Gracias a la interacción entre estructuras arbóreas, entre los diferentes términos jurídicos, entre los documentos e incluso dentro de la interacción de cada documento fuente, en su individualidad, es posible materializar el tratamiento jurídico del sistema complejo que es el ordenamiento jurídico.

Una ontología jurídica puede ser representada, utilizando la dinámica que la caracteriza al ser un sistema complejo; en un futuro próximo, la idea de evitar el análisis humano, para la construcción de la ontología, va a constituirse como uno de los focos de atención de la informática jurídica, en búsqueda de la imparcialidad y la exactitud.

En este punto, la crítica se centra en afirmar que la exactitud del análisis humano, tiene límites, de manera que no es posible afirmar que un experto en derecho, tenga la capacidad de interpretar, en una probabilidad del 100%, el significado de un concepto, de acuerdo con el contexto de tiempo y contenido, teniendo en cuenta que la sobre información, de la cual somos objeto, afecta la capacidad de lectura de toda

la información e igualmente el desarrollo de la capacidad de interpretación jurídica, que dicho sea de paso, debe ser de carácter dinámico.

La complejidad del derecho tiene fundamentos objetivos que se pueden medir, de manera que pueden ser representados dentro de una ontología, gracias a la interconexión de estructuras y al uso de teorías como la de los gráficos y estadísticas.

3.8. Hacia una concepción de las estructuras arbóreas

La gran mayoría de ontologías se construyen con estructuras arborescentes y la creación de una lista cerrada de relaciones, entre los términos que conforman la estructura arbórea. Es importante mencionar que las estructuras arborescentes, en su conjunto forman un tesoro, tema ya explicado en el primer capítulo de esta tesis.

El crecimiento del ordenamiento jurídico nacional e internacional y la rapidez de sus cambios hacen que el tratamiento de los datos de una base jurídica se torne más complejo. El impacto de este fenómeno de complejidad, dentro de los contenidos y las jerarquías jurídicas, genera la incomprensión del derecho por parte de los ciudadanos. En diciembre de 2005, el Consejo Constitucional Francés, máxima instancia judicial, sostuvo que una norma puede ser considerada inconstitucional si es inteligible y si su accesibilidad sobrepasa una “razonable complejidad”⁵⁷³.

⁵⁷³ CONSEJO CONSTITUCIONAL DE FRANCIA. Decisión, No.2005-530 (Diciembre 29, 2005). Consultada en febrero 3, 2015. <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000632146>

Con el fin de organizar las leyes a nivel europeo, se tomó la decisión de promover la codificación, catalogar todas las normas en un proyecto global que permita organizar el conocimiento jurídico bajo estructuras de títulos, capítulos, artículos y párrafos, lo que facilita la ubicación de las mismas. Así por ejemplo, en el ordenamiento jurídico francés se considera que el 75% del cuerpo normativo se encuentra en la actualidad codificado⁵⁷⁴.

Una vez culminada la primera etapa de elaboración de una ontología, dentro de la cual se elabora el tesoro, que dicho sea de paso, es una etapa que requiere la aplicación de la hermenéutica jurídica, se procede a conectar la estructura arbórea y las relaciones entre sus componentes, con el uso de un software. A pesar de las mejoras notables que se perciben, con el uso del lenguaje natural, la representación de la ontología por éste software reduce de manera inadecuada la complejidad del cuerpo jurídico⁵⁷⁵.

Para solucionar el anterior inconveniente y la complejidad del derecho, la teoría de gráficos y estadísticas⁵⁷⁶, se ha venido presentando como una solución, toda vez que permite concebir una nueva forma de representar el ordenamiento jurídico. En la arquitectura del derecho es posible observar varios niveles de organización y conexión con otros corpus jurídicos, de ahí que esta teoría considere que cada una de las normas jurídicas se encuentra cuidadosamente articuladas y encasilladas dentro de una jerarquía que la introduce en una red articulada de normas, tal y como es

⁵⁷⁴ BOUCIER, D., & MAZZEGA, P. (2007). *Codification law article and graphs*. In A. R. Lodder, & L. Mommers, *Legal Knowledge and Information System: Juris 2007*. p. 29. Netherlands: IOS Press.

⁵⁷⁵ MAZZEGA, P., BOUCIER, D., BOURGINE, P., & NADAH, N. op. cit.

⁵⁷⁶ BOUCIER, D., & MAZZEGA, P. op. cit.

representada en la pirámide de Hans Kelsen⁵⁷⁷. De ésta manera, la teoría de los gráficos y estadísticas, permite evidenciar una nueva aplicación de la pirámide que conforma el ordenamiento jurídico.

Con la teoría de los gráficos y estadísticas se puede concebir una nueva forma de representar el ordenamiento jurídico, reorganizando en un código y entre códigos la estructura jurídica y demostrando cómo, a través del tiempo, estas estructuras pueden evolucionar. Gracias a la teoría de gráficos y estadísticas podemos emplear los gráficos para representar y calcular algunas medidas de complejidad del derecho, en el ámbito de la codificación y para dar una luz al sistema jurídico y en sus evoluciones dinámicas⁵⁷⁸.

El ordenamiento jurídico se representa por un conjunto de redes interconectadas de normas (contenido cognitivo) y textos (documentos materiales). Salvo excepciones tales como la jurisprudencia o los conceptos emitidos por Entidades Públicas, la gran parte de los documentos jurídicos se organizan antes de su publicación con la correspondiente enumeración en artículos. El documento jurídico se representa, las más de las veces, en una organización jerárquica que especifica el nombre y la posición del tema, en concreto. Asimismo, es posible localizar diferentes normas dentro del mismo documento y realizar una interpretación diferente de las mismas, a pesar de que se encuentran contenidas dentro del mismo documento jurídico.

De esta manera, el conocimiento jurídico, a pesar de contar la gran mayoría de veces con una estructura que facilita su ubicación, es considerado como complejo, ya que

⁵⁷⁷ Kelsen, H. (2000). *Teoría pura del derecho* (4 ed.). México: Colofón.

⁵⁷⁸ Boucier, D., & Mazzege, P. op. cit.

existe un sin número de componentes que interactúan entre sí, un conocimiento parcial de las relaciones entre los componentes, un limitado sistema de predicción o evolución y una dependencia del sistema, cuando se observan los diferentes niveles de organización⁵⁷⁹.

El ordenamiento jurídico cuenta con gráficos subyacentes, estructura arbórea y redes, cuando se trata de normas codificadas. Un código es dividido en libros, cada libro en títulos, éstos títulos en capítulos y éstos últimos a su vez, se encuentran divididos en secciones, sub secciones y párrafos. El nivel más bajo dentro de la estructura es el artículo, sin embargo, es posible observar de igual manera la existencia de una organización jerárquica entre los artículos: párrafos con numeración romana, pueden reunir sub párrafos con numeración árabe. Se puede afirmar que la jerarquía consiste en diez niveles.

La representación formal, de la organización jerárquica de un código, es una estructura arbórea. Sin embargo, es posible encontrar dentro de un código citas o referencias a textos jurídicos externos; de esta manera es posible afirmar que existen estructuras sobre impuestas en la estructura misma del código y que, cada código a su vez, se encuentra conectado con otras estructuras arbóreas externas. De igual manera, los textos jurídicos no codificados, que no cuentan con esta estructura arbórea, se encuentran conectados con otros documentos que cuentan o no, con dicha estructura.

⁵⁷⁹ Ibidem.

Es posible observar que cualquier clase de jerarquía contenida en un código, puede explícitamente estar citada en un artículo, tal y como se puede representar mediante el siguiente gráfico, en el cual, además, se puede observar que a partir de la jerarquía se generan vértices en la siguiente estructura arbórea:

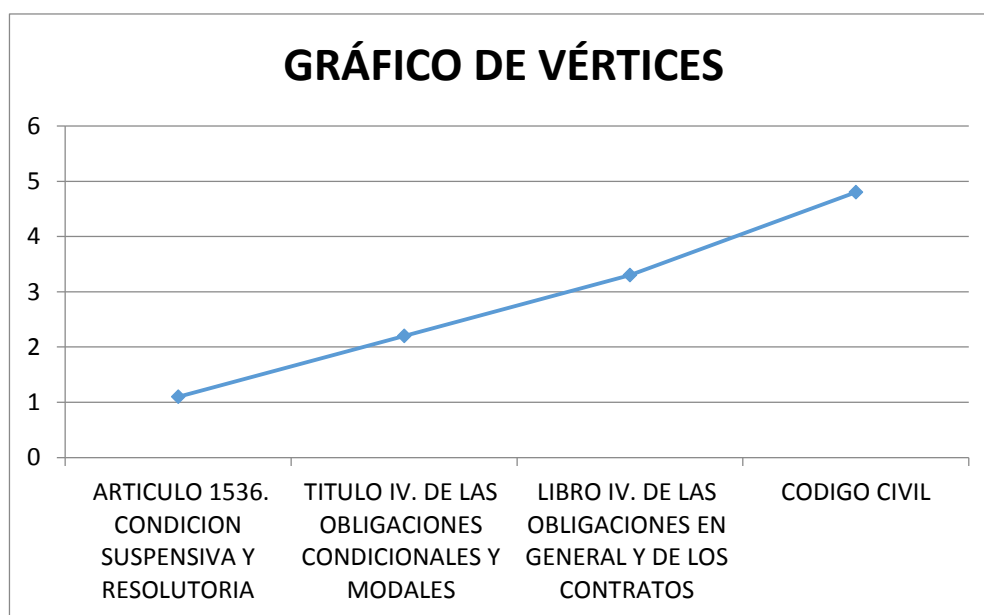
Código civil

Libro IV. De las obligaciones en general y de los contratos

Título IV. De las obligaciones condicionales y modales

Artículo 1536. Condición suspensiva y resolutoria

Figura 8. Estructura arbórea de artículos Código Civil Colombiano.



Fuente: elaboración propia con base en conceptos de estructura arbórea.

Obsérvese bien que, explícitamente, el artículo puede citar contenidos de toda jerarquía.

En el objeto anterior, el artículo 1536 del código Civil Colombiano dispone que: “La condición se llama suspensiva si, mientras no se cumple, suspende la adquisición de un derecho; y resolutoria, cuando por su cumplimiento se extingue un derecho”⁵⁸⁰

Ahora bien, basta con examinar la ubicación para poder concluir que esta disposición se aplica a todas las obligaciones en general y de manera particular, a los contratos que se encuentran sometidos al régimen civil. A pesar que no todos los niveles se encuentran necesariamente representados por cada parte individual, cualquiera de estos contenidos, si se encuentra involucrado o relacionado con otro generando un vértice.

La gráfica anterior entra en relación con otras cuando existen temas que convergen entre sí. Pensemos, por ejemplo, en la promesa unilateral de venta, que se encuentra dentro del derecho de contratos y puede contener una o varias condiciones suspensivas, para que este contrato preliminar cumpla con las condiciones de validez. El artículo 1611 del Código Civil Colombiano, subrogado por artículo 89 de la ley 153 de 1887, dispone:

“ARTICULO 1611. REQUISITOS DE LA PROMESA. La promesa de celebrar un contrato no produce obligación alguna, salvo que concurran las circunstancias siguientes:

⁵⁸⁰ Código Civil Colombiano. (1990). Bogotá: Temis. p. 352.

1a.) Que la promesa conste por escrito.

2a.) Que el contrato a que la promesa se refiere no sea de aquellos que las leyes declaran ineficaces por no concurrir los requisitos que establece el artículo 1511, 1502 del Código Civil.

3a.) Que la promesa contenga un plazo o condición que fije la época en que ha de celebrarse el contrato.

4a.) Que se determine de tal suerte el contrato, que para perfeccionarlo solo falte la tradición de la cosa o las formalidades legales.

Los términos de un contrato prometido, solo se aplicarán a la materia sobre que se ha contratado”⁵⁸¹.

Obsérvese que el artículo 1611 del Código Civil Colombiano, contiene los requisitos de la promesa en los cuales nacería una obligación, en su numeral 3 exige que “*la promesa contenga un plazo o condición que fije la época en que ha de celebrarse el contrato*”. Siguiendo la estructura arborescente el artículo se encuentra ubicado en el título XII, sobre el efecto de las obligaciones y en el libro IV de las obligaciones en general y de los contratos. En consecuencia, su estructura puede ser representada de la siguiente forma:

⁵⁸¹ *Ibíd.* p. 370.

Código Civil

Libro IV de las obligaciones en general y de los contratos

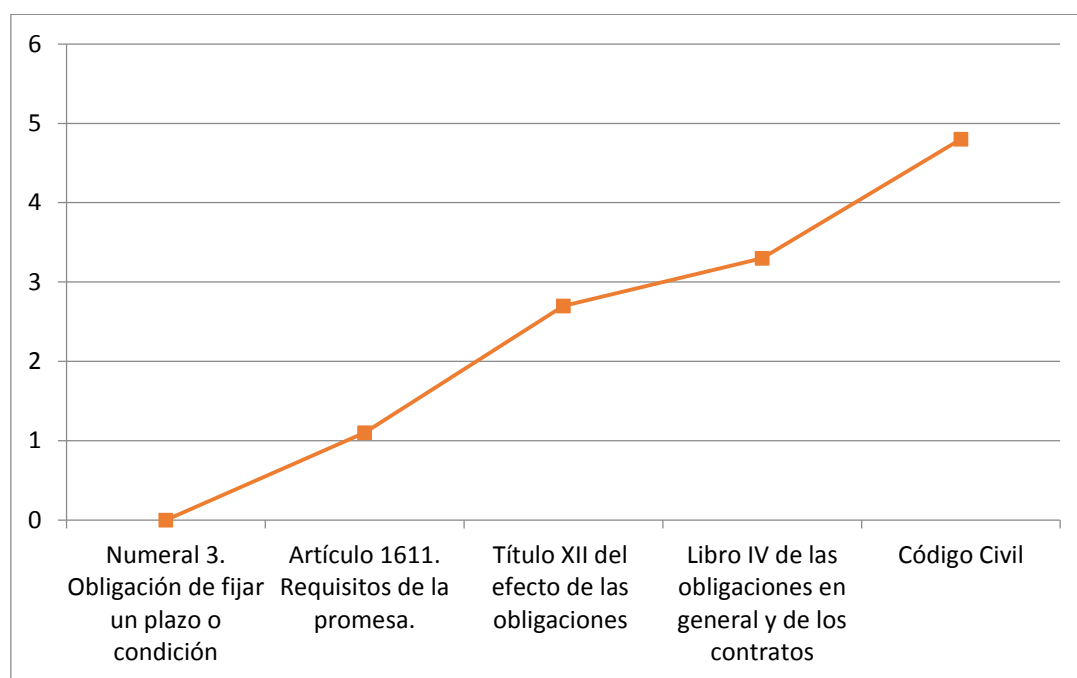
Título XII del efecto de las obligaciones

Artículo 1611. Requisitos de la promesa.

Numeral 3. Obligación de fijar un plazo o condición

El gráfico de esta estructura puede ser representado así:

Figura 9. Estructura arbórea de apartes del Código Civil Colombiano.

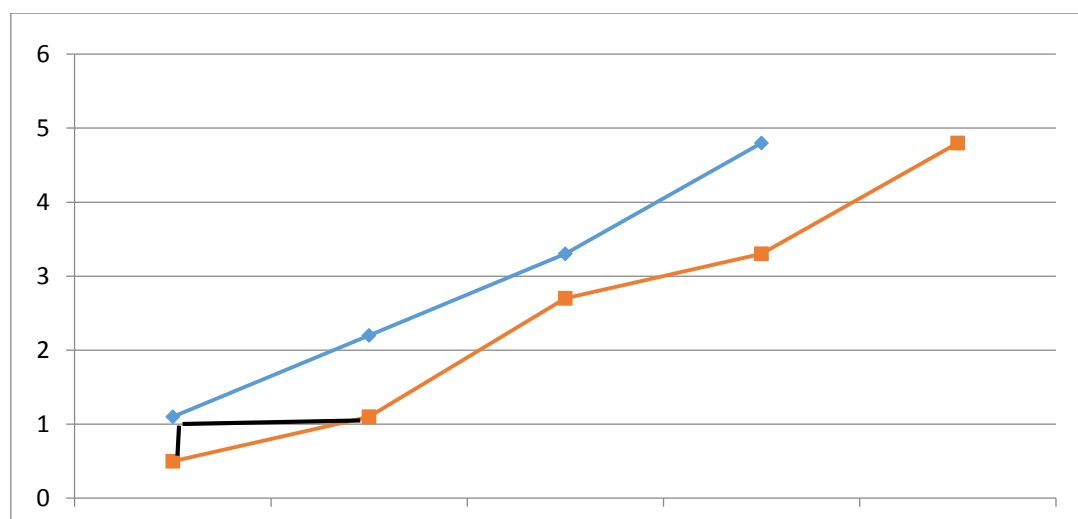


Fuente: elaboración propia con base en conceptos de estructura arbórea.

Realizando un análisis conjunto entre los dos gráficos anteriores, podemos observar que existen puntos de convergencia entre los dos, toda vez que la promesa al ser un contrato preliminar debe fijar un plazo o una condición, de acuerdo con el numeral

3°. del artículo 1611 del Código civil. Esta condición se encuentra, a su vez, definida por el artículo 1536 del código civil que define las dos clases de condiciones (suspensiva y resolutoria), por lo cual, al realizar un razonamiento jurídico estos dos conceptos deberán ir concatenados. Un gráfico que representa la creación de puntos de enlaces entre los dos artículos es el siguiente:

Figura 10. Análisis conjunto de convergencia de estructuras arbóreas.



Fuente: elaboración propia, resultado del análisis conjunto de convergencia de estructuras arbóreas, basados en los cuadros señalados en las figuras 8 y 9.

Las líneas negras del gráfico, representan la relación existente entre el artículo 1536 del Código civil, que define lo que para efectos jurídicos debe entenderse por condición suspensiva y resolutoria y, el artículo 1611 del mismo Código que señala los requisitos de la promesa y que particularmente en su numeral 3°. Consagra la obligación de fijar un plazo o condición.

En la práctica, por ejemplo, en el caso en que se pretenda realizar un contrato preliminar de promesa de venta, dentro del raciocinio jurídico que este implica, es indispensable conocer los requisitos de existencia y validez de la promesa; particularmente hay que fijar la clase de condición que va ser contenida en el contrato de promesa. Piense usted, en la inclusión de una condición suspensiva que considera válida la promesa solo en el caso en que, el beneficiario de la venta, adquiera un crédito frente a un establecimiento bancario, para pagar el precio del bien objeto del contrato.

Este raciocinio muestra como un mismo tesoro puede superponerse, tal y como se refleja en la gráfica, de manera que existen puntos de enlace que generan nuevas relaciones. El artículo 1611 numeral 3° cuando habla de la obligación de fijar una condición, hace referencia a los tipos de condiciones existentes dentro del código civil, por lo cual la palabra “condición”, en un banco de datos jurídico debe incluir la semántica de la expresión.

Tabla 1. Puntos de enlace – generación de nuevas relaciones

<p>Código Civil</p> <p>Libro IV de las obligaciones en general y de los contratos</p> <p>Título XII del efecto de las obligaciones</p> <p>Artículo 1611. Requisitos de la</p>	<p>Código civil</p> <p>Libro IV. De las obligaciones en general y de los contratos</p> <p>Título IV. De las obligaciones condicionales y modales</p> <p><u>Artículo 1536. Condición</u></p>
---	--

promesa. Numeral 3. Obligación de fijar un plazo o <u>condición</u>	<u>suspensiva y resolutoria</u>
--	--

Fuente: elaboración propia, ejemplos tomados del Código Civil Colombiano, artículos 1611 y 1536 para mostrar como el punto de enlace *condición*, genera nuevas relaciones.

Existen varios artículos que regulan “la condición” dentro del código civil, lo que confirma, una vez más, que el ordenamiento jurídico es un sistema complejo y que no basta con organizar la información, ya que es necesario relacionarla entre sí. Sobre este punto las posibilidades de enlaces de información son innumerables, el número de variables equivale al número de raciocinios jurídicos posibles y presenta una dificultad que va mucho más allá del ejemplo anteriormente expuesto.

La construcción plana de un tesoro nos facilita la tarea de organizar la información; sin embargo, la búsqueda se puede ver optimizada si dentro de la base de datos se enlazan los temas que tienen relación, lo que permitirá que los datos hablen otorgándoles una mayor semántica y ofreciendo la información más completa al usuario.

Es pertinente anotar que, el anterior ejemplo, parte de un proceso de codificación previo y analiza la estructura arbórea del Código Civil; no obstante, es necesario

tener presente que esta no es una constante dentro del ordenamiento jurídico y que existen leyes aisladas que no han sido sometidas a un proceso de codificación, pero que se encuentran vigentes y tienen la misma fuerza obligatoria que una norma contenida en un código. De manera que, a la hora de emprender un raciocinio jurídico, siempre habrá que observar los posibles enlaces con otras normas que se encuentren fuera del documento jurídico.

En este punto, las bases de datos jurídicas podrían cumplir un papel fundamental, concatenando la información, bien que esta haya o no sido codificada, siempre es posible crear nuevas estructuras arbóreas para organizar la información y relacionarlas con otras estructuras arbóreas, disímiles. Verbigracia: si se concluye un contrato de venta a distancia y se pretende saber qué obligación de información pesa sobre el vendedor, el primer reflejo consistirá en examinar las disposiciones del código civil, salvo que, si este contrato a distancia ha sido concluido entre un profesional y un consumidor, habrá que tener presente, igualmente, dentro del análisis, las normas contenidas en el estatuto del consumidor (Ley 1480 de 2011). De igual forma, si se tratase de un contrato concluido por medios electrónicos, habrá que examinar las obligaciones particulares de información existentes en materia de comercio electrónico (Ley 527 de 1999). Todo lo anterior, deberá encontrarse enlazado dentro de un banco de datos jurídicos, con el fin de examinar, desde todos los ámbitos, la regulación del tema objeto de la búsqueda.

Para concluir el planteamiento expuesto anteriormente, existen diferentes tipos de vínculos que se pueden presentar: el primero, en la relación misma que puede tener un artículo dentro de su estructura arbórea en términos de jerarquía; el segundo, en

los vínculos existentes dentro de una misma estructura arbórea, sea esta codificada o no; aquí será posible crear enlaces entre sus descriptores con el fin de relacionar la información, y en tercer lugar, es posible relacionar las diferentes estructuras arbóreas sea que estas provengan de un proceso de codificación o no.

Estos tres tipos de vínculos otorgan la posibilidad de introducir inteligencia, a los datos, toda vez que se enlaza el conocimiento de varias fuentes del derecho, lo que además de facilitar la búsqueda, otorga una visión más amplia y transversal, indispensable en el ordenamiento jurídico, ya que el raciocinio jurídico es producto de la comprensión de un sistema cruzado. Verbigracia: un jurista que elabora la fusión de dos sociedades, no puede pretermir la aplicación de las leyes tributarias que regulan la materia; y un usuario, de un banco de datos jurídico quedaría parcialmente informado si al realizar una búsqueda sobre “la oferta” no obtuviese información sobre “la retractación”.

Ahora bien, un artículo que se refiere por ejemplo, a otro artículo integrado en un contexto jurídico diferente, a causa de su contenido genérico, puede hacer alusión al primer artículo. A esto se le ha denominado enlace de segunda clase. Por ejemplo, un artículo que regula el manejo del agua puede hacer referencia a una regla que regula la responsabilidad de una autoridad territorial. Esta clase de enlaces, expresa una “selección legal”, orientada de los nodos a los enlaces. El inicio de los vértices es siempre un objeto de la clasificación del artículo y puede señalar si tiene una simple o compleja subdivisión. El vértice final puede ser un objeto de cualquier clase,

perteneciente al código de medio ambiente, a otro código, o a un texto jurídico de derecho interno o a un cuerpo jurídico del derecho europeo o internacional⁵⁸².

Es evidente la complejidad de la estructura jurídica del ordenamiento; sin embargo, el hecho de crear jerarquías permite establecer relaciones entre otras fuentes jurídicas para conectarlas entre sí. Una de las formas de observar el contenido del ordenamiento jurídico, es por conducto de la ontología, entendida como un conjunto de términos que se encuentran enlazados a otros términos, de manera que esas redes forman cuencas semánticas. En el futuro, diversos cuestionamientos jurídicos de alta complejidad en derecho, podrán ser útilmente abordados mediante una representación gráfica⁵⁸³.

Es importante precisar que, si bien examinar la estructura arborescente de un código o una ley permite sin duda contextualizar, se hace necesario conectar esta estructura compleja con su contenido concreto, es decir, con sus reglas y particularmente con el significado que cada una de ellas contienen. En este punto, a diferencia de la creación de un buscador general, un banco de datos jurídicos deberá darle prelación al significado semántico de cada palabra, en cada regla particular, toda vez que en derecho el uso del lenguaje natural podría desorientar el sentido del contenido de un documento jurídico. La hermenéutica juega, igualmente, un papel fundamental en la creación de una ontología jurídica.

⁵⁸² BOUCIER, D., & MAZZEGA, P. op. cit.

⁵⁸³ *Ibíd.*

3.9. La aplicación de la probabilidad a los conceptos ontológicos

Los contenidos del documento jurídico pueden ser medidos con el uso de la estadística y la aplicación de fórmulas, esto con la finalidad de observar la interacción entre los términos, el conjunto de términos y los contextos en los cuales son ubicados dentro de un banco de datos jurídicos y en cada documento individual. Gracias a estas herramientas, es posible analizar y observar cómo evolucionan las relaciones entre los conceptos, aplicando la probabilidad.

Para explicar este proceso es necesario situarse dentro de un contexto particular. Piense usted en un conjunto de leyes (fuentes formales) que se han recopilado para la creación de la ontología, concretamente, bajo un proceso de codificación. Los códigos cuentan con una estructura intrínseca, arbórea, en la cual existen subdivisiones en partes, libros, títulos, capítulos, secciones, párrafos, artículos e incisos. Dentro de este cuerpo es posible conocer en cuantas ocasiones se ha hecho referencia, a un término jurídico específico.

Para conocer el número de veces en que se ha citado un término jurídico, dentro de un código, que se pretende incluir en un banco de datos, es posible aplicar una fórmula para establecer la probabilidad.

En la informática jurídica del pasado, no se aplicaban los cálculos de probabilidad, su empleo no se opone a los métodos tradicionales de tratamiento de la información; por el contrario, se presenta como una herramienta suplementaria, que en términos de pertinencia arroja resultados que permiten de una manera u otra remplazar el

raciocinio jurídico humano y automatizar las relaciones entre términos jurídicos, de manera que se crean enlaces que permiten, incluso, conocer las repercusiones de un cambio de término, dentro del ordenamiento jurídico.

La posibilidad de establecer relaciones entre los diferentes conceptos, crea relaciones implícitas y explícitas entre los mismos. El suceso de la aplicación práctica de las probabilidades, si bien no permite automatizar la construcción de las ontologías jurídicas, si permite evidenciar y comprender la complejidad del sistema jurídico, enriqueciendo el tratamiento y la búsqueda de la información toda vez que la lectura de las probabilidades puede aportar más información que una lectura semántica, al conectar los conceptos entre sí.

A pesar que en un futuro próximo, no será posible prescindir del análisis humano como base del tratamiento de una ontología, la aplicación de las probabilidades permite capturar la complejidad del contenido. La aplicación de la lógica y de la estadística no son herramientas excluyentes; es posible, incluso, aplicar la probabilidad entre dos o más ontologías jurídicas distintas, con el fin de compararlas, mejorarlas, complementarlas, fusionarlas o analizarlas conjuntamente.

3.10. Formalización de la ontología

Seguidamente, la ontología es formalizada con la ayuda de un editor. Dentro de los editores en particular, existe uno denominado *Protégé*⁵⁸⁴, que “es un editor de

⁵⁸⁴ STANFORD CENTER FOR BIOMEDICAL INFORMATICS RESEARCH. (2015). Protégé. Consultada en febrero 3, 2015. <http://protege.stanford.edu/>

ontologías distribuido por la Universidad en informática médica de Stanford. Protégé no es una herramienta únicamente dedicada a OWL, sino un editor altamente extensible, capaz de manipular formatos muy diversos (...) Protégé es una herramienta empleada por los desarrolladores y los expertos en el tema para desarrollar sistemas basados en los conocimientos (Ontologías). Las aplicaciones desarrolladas con Protégé son empleadas en la resolución de problemas y la toma de decisiones en un tema particular. Protégé es igualmente una plataforma extensible, gracias a un sistema de plug-ins⁵⁸⁵ que permite generar contenidos multimedia, interrogar, evaluar y fusionar las ontologías, etc.”⁵⁸⁶. Es decir que, protégé, es un ejemplo de aplicación de la inteligencia artificial.

Para materializar la ontología, el editor, especialmente *Protégé*, “posee una interface que le permite manipular todos los elementos de una ontología: clase, meta clase, propiedad, instancia, etc.”⁵⁸⁷. Este editor es versátil, pues es posible hacer uso del mismo, en cualquier ámbito del conocimiento, incluyendo el jurídico, en el cual es necesario organizar jerárquicamente la información. Otra ventaja de este editor radica en la posibilidad de hacer uso de los diferentes lenguajes informáticos, utilizados para crear la ontología (OWL, DAML, OIL, etc.).

Una de las críticas que se le puede hacer, a este editor, radica en que este carece de un campo específico para introducir la información lingüística, es decir, la definición de términos; sin embargo, es posible exportar del DOE (Differential Ontology

⁵⁸⁵ Un Plug-in, añadido o complemento es una aplicación que aporta una función nueva y específica a otra aplicación o programa, esta se ejecuta con el uso de la aplicación principal.

⁵⁸⁶ BENHMIDI, H., & MAZOUZI, A. (2011). Construction et manipulation d'une ontologie médicale. Consultada en marzo 2, 2015. <http://dSPACE.univ-tlemcen.dz/bitstream/112/1062/9/Annexe-Protege.pdf>

⁵⁸⁷ Ibídem.

Editor) hacia Protégé, para permitir al usuario la conservación de la información lingüística durante el proceso de construcción de la ontología⁵⁸⁸.

Para formalizar y crear la ontología con Protégé, es necesario, en un primer momento, diseñar un proyecto que, técnicamente para su creación, es obligatorio elegir el lenguaje o formato; en segundo lugar, crear las clasificaciones; en tercer lugar, crear las propiedades o slots y, finalmente, crear instancias o sub clasificaciones.

Una vez la ontología se encuentra creada, cuando se surte el procedimiento de formalización con el uso del editor, es posible integrar las nociones lingüísticas, lo que permite introducir la noción semántica asociada al concepto que se encuentra en la ontología.

3.11. La materialización semántica de la ontología

Las nociones lingüísticas constituyen un punto neurálgico dentro de una ontología, toda vez que proporcionan la información fundamental de la semántica de los conceptos. Para lograr el objetivo de crear contenidos semánticos, existen plataformas que reúnen toda la cadena de construcción de una ontología, es decir, desde la extracción de términos hasta la formalización de la ontología, gracias al uso de un editor y, una vez reunida la cadena, la plataforma otorga los vínculos entre los conceptos y la ontología. No obstante, los vínculos que se construyen entre los conceptos y la ontología deben ser almacenados manualmente. En este punto, el

⁵⁸⁸ MAZZEGA, P., BOUCIER, D., BOURGINE, P., & NADAH, N. op. cit.

lenguaje natural juega un papel preponderante para la buena culminación de la ontología.

Dentro de estas plataformas, existe una en particular denominada DAFOE, que es una de uso abierto, para crear ontologías a partir de textos. Ella tiene por objetivo organizar grandes volúmenes de información. La concepción de esta plataforma ha sido realizada dentro del proyecto denominado ANR Dafoe4App. A pesar de los avances logrados, el proyecto no ha podido llegar a su conclusión, por falta de financiamiento.

Con la creación de la plataforma multimétodos y multimodelos DAFOE, se busca, principalmente, construir las ontologías con un procedimiento unificado, toda vez que esta ofrece varios recursos, a saber: “(1) las redes terminológicas se organizan durante el análisis de datos, (2) un nivel termino-conceptual en el cual los conceptos son organizados y (3) un nivel en el cual la ontología es formalizada”⁵⁸⁹. De esta manera, es posible encontrar dentro de un solo programa la unificación del proceso automatizado de creación de la ontología.

El procedimiento, dentro de la plataforma DAFOE, pretende ofrecer un conjunto de herramientas y otorgar, en un primer tiempo, la posibilidad de visualizar los resultados de las herramientas TAL (Tratamiento Automático del Lenguaje). Posteriormente, dar la posibilidad al usuario de realizar un estudio lingüístico o terminológico. En un tercer tiempo, crear una red término-conceptual o término-

⁵⁸⁹ SZULMAN, S., HERNÁNDEZ, N., TEGUIAK, V., CHARLET, J., NADAH, N., & OTROS. (2010). *DAFOE: une plateforme multiméthodes et multimodèles pour*. *Congres Francophone AFRIF-AFIA Reconnaissance des Formes et Intelligence Artificielle*. l'Université de Caen Basse-Normandie. Consultada en marzo 28, 2015. <ftp://ftp.irit.fr/IRIT/IC3/demoRFIA.pdf>

ontológica y concluir con la formalización y creación de una ontología, con exportación en lenguaje OWL.

De otra parte, existen otras plataformas que ya se encuentran finalizadas y que contribuyen a materializar semánticamente la ontología; tal es el caso de TERMINAE⁵⁹⁰, que es una plataforma que contribuye a la creación de los recursos término-ontológicos, a partir de los textos, permitiendo la materialización conceptual de los documentos, objeto de análisis. Es posible, dentro de ella, construir tesauros, ontologías y crear recursos termino-ontológicos, con bases lingüísticas.

La plataforma es realizada por la aplicación eclipse y puede ser utilizada con el lenguaje OWL y SKOS. No obstante, el software permite, únicamente, que los documentos base o documentos fuente para crear la ontología, se encuentren redactados en inglés o francés, por lo cual para efectos de crear una ontología jurídica en español, el programa no resultaría adaptado.

Una crítica que se le puede realizar, a estas herramientas, es que ellas no tratan los vínculos implícitos que no se encuentran señalados por el analista de datos y de igual forma estos programas no tienen presente la dinámica del ordenamiento jurídico para adaptar en el tiempo la ontología⁵⁹¹.

⁵⁹⁰ Sitio Web del proyecto TERMINAE <http://lipn.univ-paris13.fr/terminae>

⁵⁹¹ MAZZEGA, P., BOUCIER, D., BOURGINE, P., & NADAH, N. op. cit.

CONCLUSIONES

Para concluir, sobre lo hasta aquí desarrollado, podemos mencionar que, desde los años 1990, se ha venido presenciando el movimiento por la modelización del conocimiento, que emerge apoyado de la noción de ontología. La informática jurídica no fue ajena a este fenómeno. Es importante precisar que la construcción de una ontología, a partir de los documentos jurídicos, ha sido objeto de estudio, no solamente en el ámbito de la informática jurídica sino también, de la ingeniería de ontologías.

La definición de un método, para la construcción de una ontología jurídica, resulta fundamental, cualquiera que sea el modelo metodológico que se pretenda implementar. Este deberá contener como mínimo cuatro etapas, a saber: la primera, la construcción de un cuerpo ontológico jurídico (entiéndase por esto la recopilación, clasificación organización de documentos); en segundo lugar, se debe realizar un análisis lingüístico de los documentos jurídicos; luego, es necesario realizar la conceptualización de los términos y finalmente, es necesario materializar o formalizar la ontología.

En la actualidad, a pesar de la existencia de varias plataformas de tratamiento automático del lenguaje, que analizan de forma automática los documentos, desde el punto de vista sintáctico y estadístico, en la informática jurídica no se ha logrado establecer ningún procedimiento de creación de ontologías, de aplicación general. De igual forma, no existen herramientas informáticas unificadas que permitan establecer la trazabilidad, la concepción progresiva y secuencial de una ontología jurídica. Sin

embargo, es posible encontrar investigaciones y proyectos ya materializados que se encaminan, a crear plataformas de acceso libre, para facilitar la tarea de materializar y formalizar las ontologías. Tal es el caso del editor Protégé y de la plataforma DAFOE.

La concepción de una ontología jurídica que pueda ser susceptible de ser desarrollada, dentro de un programa unificado, creado para tales efectos y aceptado de manera generalizada, en el campo de la informática jurídica, presenta dificultades en la actualidad.

Para remplazar la ausencia de un programa unificado que permita la creación de una ontología jurídica, se han creado métodos separados que en su conjunto confluyen a la culminación de la ontología. Estos métodos pueden ser divididos en dos partes, a saber: los métodos materiales de realización de una ontología jurídica, y, los métodos formales de la implementación de la ontología.

Dentro de los métodos materiales de realización de una ontología jurídica, encontramos los siguientes:

1. El análisis del contenido de los documentos.
2. La creación de los tesauros o estructuras arbóreas y su interconexión. Dentro de este punto se incluye la teoría de gráficos y estadísticas, la interconexión entre tesauros y ramas del derecho, independientes, que en puntos determinados puedan encontrarse relacionadas con temas jurídicos concretos, lo que proporciona una

visión más amplia del tema, a la hora de realizar una búsqueda dentro de un banco de datos jurídico.

3. El uso del contenido de los códigos, que bien pueden proporcionar estructuras arborescentes o fragmentos de estructuras arborescentes. Este apoyo en la codificación puede orientar cada una de las bases de datos jurídicas, en el caso en que su corpus jurídico haya sido objeto de unificación. Lo cual, en términos hermenéuticos, se traduce en la interpretación de una fuente formal, aceptada en un consenso jurídico y socialmente reconocido como válido.

A su vez, dentro de los métodos formales de la implementación de la ontología, se encuentran las herramientas informáticas que sirven de apoyo para materializar formalmente la ontología jurídica. Los programas existentes en la actualidad son, entre otros, los siguientes:

- ✓ el programa Protégé de la Universidad de Standford, el cual permite representar formalmente una ontología.
- ✓ Las plataformas de Tratamiento Automático de Lengua (TAL), que otorgan la posibilidad de realizar un tratamiento automatizado de los documentos jurídicos y de crear notas desde el punto de vista estadístico y sintáctico, de cada documento analizado.
- ✓ La plataforma DAFOE, que es un proyecto creado en Francia con la iniciativa de varias universidades y en CNRS de Paris.

Sin embargo, como se ha mencionado en líneas anteriores, no existe un consenso actual sobre la metodología y el procedimiento generalmente aceptado para la creación de una ontología jurídica.

Tal vez, en un futuro, será posible presenciar el avance de la informática jurídica hacia la creación de un sistema inteligente, capaz de procesar y simular el razonamiento jurídico, otorgando respuestas que conecten los conceptos jurídicos entre sí y haciendo uso de la hermenéutica, para facilitar la aplicación del ordenamiento jurídico. De igual forma, será posible en un futuro próximo, crear bases de datos jurídicas casuísticas que permitan orientar las decisiones de los jueces, gracias a la aplicación de la estadística y al apoyo de máquinas inteligentes que puedan razonar jurídicamente.

En un futuro, un método genérico y automático se posicionará para transformar, cualquier banco de datos jurídicos u ordenamiento. En este punto, la web *eneactif* juega un papel preponderante, toda vez que esta permite independencia y auto representación del contenido. De manera que se va a requerir de intervención, para realizar la interacción entre la información y el usuario.

El término “*eneactif*”, en derecho, hace referencia a la forma como un sistema de coincidencias puede contribuir con sus acciones a responder a las exigencias jurídicas, de una situación concreta. Lo que, en otras palabras, se concretaría en la interacción entre el sistema (software, computador, máquina) y el usuario. De

manera que, las relaciones hombre-máquina, serían directas y de la misma manera en que un individuo podría realizarlo.

Esta es la nueva perspectiva de la informática jurídica que se encuentra fundada sobre los ya estudiados sistemas complejos; aquí el concepto de *web enactif*, juega un papel fundamental en el campo de las ontologías jurídicas; el estudio y aplicación de este concepto podrá, en un futuro, crear la autonomía de las bases de datos, contando con la participación de expertos.

“En las próximas etapas, un método genérico automático para transformar cualquier cuerpo jurídico en web enactif, surgiendo de forma continua con la actividad de expertos y de comunidades de usuarios podrá ser aplicado. La web enactif significa que ella puede autorepresentarse o “tener rendimiento”; es una forma de conocimiento autopoietico, que ciertos juristas como Luhmann habían ya identificado antes de la introducción de la informática. Por ejemplo, cualquier configuración en un sub conjunto jurídico será capaz de producir un nuevo mapa. Será posible construir diccionarios “bottom-up” y ontologías continuas producidas por la plataforma susceptible de medir la tasa de evolución y de coevolución de la ley, la jurisprudencia, la doctrina, los diccionarios y tesauros especializados al interior de un Estado, o entre los diferentes Estados europeos. Esta opción podría ayudar al legislador europeo a comparar (casi en línea) la evolución de las ontologías jurídicas en Europa”⁵⁹².

⁵⁹² MAZZEGA, P., BOUCIER, D., BOURGINE, P., & NADAH, N. op. cit.

Los avances de la informática jurídica, desde sus inicios, son fácilmente constatados; las bases de datos actuales así como las de un futuro próximo, no se conformarán con el simple objetivo de organizar la información, para facilitar la búsqueda como si se tratara de un fichero o de una biblioteca; estas bases de datos pretenden lograr la interconexión de los documentos y conceptos en ella contenidos, para entregar resultados al usuario, con contenidos semánticos que conecten los conceptos, relacionándolos, para facilitar su aplicación.

Nos encontramos frente a retos mayores. La inteligencia jurídica artificial puede contribuir a la actividad del operador jurídico (magistrados, jueces, abogados, notarios, etc.). Los resultados semánticos y racionales pueden facilitar la toma de decisiones, la aplicación del derecho atendiendo a los cambios que se presentan, dentro del ordenamiento jurídico en tiempo real.

No obstante, es importante precisar que la inteligencia jurídica artificial se encuentra lejos de remplazar la labor que realiza un operador jurídico y que, este concepto de “inteligencia”, al ser materializado, no podría alcanzar totalmente la forma, como un jurista realiza el razonamiento, cuando tiene en sus manos un conjunto de conceptos y de documentos jurídicos, en especial, si se tiene en cuenta que la experiencia juega un papel fundamental.

La inteligencia jurídica artificial debe cumplir un rol de acompañamiento del jurista de la actualidad, que se encuentra ejerciendo dentro de la sociedad de la información dinámica, que día a día, sufre cambios importantes dentro del ordenamiento jurídico, a causa del desarrollo de áreas del derecho, aún más recientes y que deben responder

y adaptarse a los nuevos supuestos de hecho que se presentan en la sociedad e igualmente a causa de los avances jurídicos que surgen dentro de las áreas del derecho clásico.

Las bases de datos jurídicas especializadas, en cada una de las ramas del derecho, tendrán la nueva tarea de reunir, en una sola base de datos, todo el conjunto de fuentes jurídicas, representando así al sistema complejo del ordenamiento jurídico unificado, toda vez que la interdependencia de las áreas del derecho es evidente, de lo contrario se corre el riesgo de obtener resultados parcializados, al emprender una búsqueda.

Las bases de datos jurídicas y sus tesauros, deben actualmente ir encaminadas a su unificación, con el fin de aprovechar la aplicación de las herramientas que ofrece una ontología jurídica, que tiene como razón de existencia, la conexión y el análisis del lenguaje natural, dentro de un sistema complejo.

BIBLIOGRAFÍA

- Agencia Central de Inteligencia. CIA. (2013). *The World Factbook*. Washington, DC: Central Intelligence Agency. <https://www.cia.gov/library/publications/the-world-factbook/rankorder/2153rank.html?countryName=Colombia&countryCode=co®ionCode=sa&rank=21#co>
- Agudelo Benjumea, M. M. (s.f). Los Metadatos. *Gestión de Contenidos de Educación Virtual de Calidad*. Ministerio de Educación Nacional de Colombia y Universidad de Antioquia. Consultada en febrero 12, 2015.http://aprendeonline.udea.edu.co/lms/men/docsoac3/0301_metadatos.pdf
- Alfonso Moreira, C. G. (2002). Las bases de datos de la UE. *Revista Valenciana de Economía y Hacienda*, 249-262. Consultada en mayo 31, 2012. http://portales.gva.es/c_economia/web/rveh/pdfs/n5/materiales5.pdf
- Alvite Díez, M. L. (2003). Tendencias en la investigación sobre recuperación de información jurídica. *Revista Española de Documentación Científica*, 26(2). doi:10.3989/redc.2003.v26.i2.137
- Alvite Díez, M. L. (2009, Enero-Junio). Las bases de datos jurídicas y el uso del lenguaje XML en España. *Scire*, 15(1), 2. <http://www.iversid.eu/ojs/index.php/scire/article/view/1768/1739>
- Alvite Díez, M. L. (2012). El uso de vocabularios controlados en sistemas de información jurídica: evolución y tendencias actuales de representación. *Revista Scire*, 18(1).

Alvite Díez, M. L., Pérez-León, B., Martínez González, M., & Blanco, V. (2010).

Propuesta de representación del Tesauro Eurovoc en SKOS para su integración en sistemas de información jurídica. *Revista Scire*, 6(2), 47-51.

Ambientalex.info. El portal ambiental. (2015). Consultada en agosto 28, 2014.

<http://www.ambientalex.info/>

Arango Vásquez, S. I., & Ricaurte Avendaño, A. (2006). *Manual de herramientas tecnológicas*. Medellín: Universidad de Medellín.

Araña Web. (2014, Octubre 21). Consultada en febrero 5, 2014.

http://es.wikipedia.org/w/index.php?title=Ara%C3%B1a_web&oldid=77670867.

Argentina Judicial. (n.d.). Consultada en julio 10, 2014.

<http://argentinajudicial.com.ar/>

Argentina. Boletín Oficial. (n.d.). Consultada en julio 6, 2014.

<http://www.boletinoficial.gov.ar/Inicio/Index.castle>

Argentina. Corte Suprema de Justicia. (n.d.). Consultada en julio 5, 2014.

<http://www.csjn.gov.ar/jurisprudencia.html>

Argentina. Ministerio de Economía y Finanzas Públicas. (2005). Consultada en julio

6, 2014. Infoleg. Información Legislativa y Documental.: http://infoleg.mecon.gov.ar/?page_id=122.

Argentina. Ministerio de Justicia y Derechos Humanos. (2015, Febrero). Consultada

en julio 6, 2014. Departamento de Biblioteca y Centro de Documentación: <http://www.biblioteca.jus.gov.ar/jurisprudencia-argentina.html>,

Argentina. Red Federal de Bibliotecas Parlamentarias. REFBIPAR. (2006).

Consultada en julio 5, 2014. <http://refbipar.bcn.gov.ar/>

Article 29 Data Protection Working Party. Opinion 03/2013. (2013, Abril 2).

http://idpc.gov.mt/dbfile.aspx/Opinion3_2013.pdf

Assemblée Nationale France. (n.d.). Consultada en marzo 13, 2015.

<http://www.assemblee-nationale.fr/index.asp>

Association of the Councils of State and Supreme Administrative Jurisdictions. ACA-

Europe. (2015). Consultada en agosto 29, 2014. Bases de Datos:

<http://www.aca-europe.eu/index.php/en/dec-nat-en>

Baader, F., Horrocks, I., & Sattler, u. (2015, Febrero 2). Description Logics as

Ontology Languages for the Semantic Web. *Mechanizing Mathematical*

Reasoning, 228-248. Consultada en febrero 2, 2015.

<http://www.sfu.ca/~mhatala/iat881/papers/baader-et-al.pdf>

Bachimont, B. (2004). *Art et sciences du numérique: Ingénierie des connaissances et*

critique de la raison computationnelle. Université de Technologie de

Compiègne .

Bauer-Bernet, H. (1982). *Legal thesaurus and data processing.* Amsterdam.

Benhmidi , H., & Mazouzi , A. (2011). *Construction et manipulation d'une ontologie*

médicale. Consultada en marzo 2, 2015. [http://dspace.univ-](http://dspace.univ-tlemcen.dz/bitstream/112/1062/9/Annexe-Protege.pdf)

[tlemcen.dz/bitstream/112/1062/9/Annexe-Protege.pdf](http://dspace.univ-tlemcen.dz/bitstream/112/1062/9/Annexe-Protege.pdf)

Benjamis, R., Breukers, J., Casanovas, & Gangemi, A. (2005). Law and the

Semantic Web, an introduction. *Law and the Semantic Web*, 1-17.

Béra, M., & Mechoulam, É. (1999). *La machine internet*. Paris: Ediciones Odile Jacob.

Berners-Lee, T., Hendler, J., & Lasilla, O. (2001). The Semantic Web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, 284(5), 34-43.

Biblioteca Andrés Aguilar Mawdsley. (2011). Consultada en junio 23, 2014. <http://aciempol.msinfo.info/>

Binde, J. (2005). Hacia las sociedades del conocimiento: Informe Mundial de la UNESCO. Paris: Unesco. <http://unesdoc.unesco.org/images/0013/001321/132114f.pdf>

Bing desplazó a Yahoo! y es el segundo buscador más usado en el mundo. (2011, Marzo 3). Enter.Co: http://www.enter.co/cultura-digital/entretenimiento/bing-desplazo-a-yahoo-y-es-el-segundo-buscador-mas-usado-en-el-mundo/?s_cid=1

Boletín Oficial del Estado. (n.d.). Consultada en enero 3, 2013. <http://www.boe.es/>

Boucier, D., & Mazzega, P. (2007). Codification law article and graphs. In A. R. Lodder, & L. Mommers, *Legal Knowledge and Information System: Juris 2007* (pp. 29-38). Netherlands: IOS Press.

Brasil. Senado. Prodasen. (n.d.). Consultada en junio 24, 2014. <http://www.senado.gov.br/senado/prodasen/>

Bray, T., & Bosak, J. (1999). XML and the second-generation web. *Scientific American*, 280(5), 89-93. <https://courses.washington.edu/infx598/win12/secondGenerationWeb.pdf>

Brecha digital. (2015, enero 17). http://es.wikipedia.org/w/index.php?title=Brecha_digital&oldid=79306248.

British and Irish Legal Information Institute. (2010, Noviembre 25). Consultada en agosto 20, 2014. Bases de Datos: <http://www.bailii.org/databases.html>

Buchanan, B., & Headrick, T. (1970, Noviembre). Some speculation about artificial intelligence and legal reasoning. *Stanford Law Review*, 23(1), 40-62. doi:10.2307/1227753

Bundesanzeiger Verlag GmbH. (n.d.). Consultada en agosto 30, 2014. <http://www.bundesanzeiger-verlag.de/ueber-uns.html>

Bundesministerium der Justiz und Fur Verbraucherschutz. (n.d.). Consultada en agosto 30, 2014. <http://www.gesetze-im-internet.de/titelsuche.html>

Cáceres Nieto, E. (1984). *Teoría del abstract legislativo*. México: Universidad Nacional Autónoma de México. UNAM.

CADE. Cuatro Décadas Brindando Servicios de Alta Calidad. (2012). Consultada en agosto 26, 2014. www.cade.com.uy/

Cafassi, E. (1998). *Internet: Políticas y comunicación*. Buenos Aires, Argentina: Biblos.

- Calvo Prieto, J. M., Rocha Freire, I., Desongles Corrales, J., & Santos Rodríguez, M. (2006). *Auxiliar de servicio de la Universidad Politécnica de Valencia* (Vol. 2). España: Editorial Mad S.L.
- Campi, C. (1982). *Intelligenza artificiale e sistema informativi giuridici*. Italia: Istituto di Teoría e Techniche dell'Informazione Giuridica. Consultada en marzo 3, 2015. http://www.ittig.cnr.it/EditoriaServizi/AttivitaEditoriale/InformaticaEDiritto/1982_02_079-091_Ciampi.pdf
- Caplan, P. (1995). You Call It Corn, We Call It Syntax-Independent Metadata for Document-Like Objects. *The Public-Access Computer Systems Review*, 6(4).
- Casanovas, P. (2012, Enero-Junio). Algunas líneas de investigación en gestión del conocimiento jurídico: web semántica, ODR y derecho racional. *Revista Scire*, 18(1), 15-28. <http://ibersid.eu/ojs/index.php/scire/article/viewFile/3943/3672>
- Castañeda Ponce, D., Hernández Martínez, M., & González Galván, J. (1989). El léxico como instrumento lingüístico del sistema UNAM-JURE. In J. Bilon , H. Fix Fierro, E. Cáceres , & S. Matute, *Diálogo sobre la informática jurídica*. México: Universidad Nacional Autónoma de México. UNAM.
- Castells, P. (2003). La Web Semántica. In C. Bravo Santos, & Á. Redondo Duque, *Sistema interactivos y colaborativos en la Web* (p. 280). Cuenca: Universidad de Castilla-La Mancha. <https://goo.gl/6CXQrO>
- Castro Fernández, J. (1992). *Juristas y computadoras*. San José de Costarica: Juriscomputación.

CEN MetaLex. Open XML Interchange Format for Legal and Legislative Resources.

(2006, Julio 7). <http://www.metalex.eu/>

Centre National de Documentation Juridique. (2012-2015). Consultada en

septiembre 1, 2014. <http://www.cndj.ci/index.php>

CIJUL en Línea. Centro de Información Jurídica en Línea de Costa Rica. (2014).

Consultada en julio 29, 2014. <http://cijulenlinea.ucr.ac.cr/>

Código Civil Colombiano. (1990). Bogotá: Temis.

Codina, L., & Pedraza Jiménez, R. (2011, Septiembre- Octubre). Thesaurus y

ontologías en sistemas de información documental. *El profesional de la*

información, 20(5), 555-563. <http://www.elprofesionaldelainformacion.com/>

contenidos/2011/septiembre/10.pdf

Colombia. Corte Constitucional. (n.d.). <http://www.corteconstitucional.gov.co/>

Commission Européenne. (2014, enero 18). <http://ec.europa.eu/legislation/>

index_fr.htm

Commission Nationale de l'Informatique et des Libertés. (n.d.). Vie privée

à l'horizon 2020 Paroles d'experts. *Cahiers Ip. Innovation & prospective*, 1-

60.

Congreso de la República de Colombia. (1993, diciembre 23). Ley 100 por la cual se

crea el sistema de seguridad social integral y se dictan otras disposiciones.

Diario Oficial, No.41.148.

Conseil constitutionnel. (n.d.). Consultada en marzo 13, 2015. Conseil constitutionnel. Les décisions: <http://www.conseil-constitutionnel.fr/conseil-constitutionnel/francais/les-decisions/les-decisions.95486.html>

Consejo de la Unión Europea. (2014, enero 18). <http://www.consilium.europa.eu/es/council-eu/>

Consiglio Nazionale della Recherche. Istituto di Teoria e Tecniche dell'Informazione Giuridica. (n.d.). Consultada en mayo 8, 2014. www.cnr.it/istituti/Descrizione.html?cds=104

Córdoba (Argentina). Codex. Sistema de Ordenamiento y Sistematización de Normas Provinciales. (2015, Enero 31). Consultada en julio 6, 2014. <http://codex.cba.gov.ar/Paginas/Busqueda.aspx>

Córdoba (Argentina). Dirección de Informática Jurídica. (n.d.). Consultada en julio 6, 2014. Legislación Provincial: <http://web2.cba.gov.ar/web/leyes.nsf/fri?OpenForm>

Cornell Law School. (1992). Consultada en septiembre 02, 2013. <https://www.law.cornell.edu/>

Cortes de Catilla y León. (n.d.). Consultada en marzo 14, 2015. <http://www.ccyl.es/export/sites/ccyl/index.html>

Cour de cassation. (n.d.). Consultada en marzo 13, 2015. Cour de cassation. Jurisprudence: https://www.courdecassation.fr/jurisprudence_2/

Cukier, K., & Mayer-Schoenberger, V. (2014). *Big Data: La révolution des données est en marche*. Robert Laffont.

Dalloz.fr. (2015). Consultada en marzo 13, 2015. Dalloz.fr: <http://www.dalloz.fr/>

Dávila, J. (2006). La lógica, los agentes y la web semántica. *Revista de la Universidad de los Andes de Merida*.

De Kunder, M. (2011, julio 12). *The size of the World Wide Web. The Internet*. (U. d. Tilburg, Ed.). <http://www.worldwidewebsize.com/>

De la Cueva, J. (2011, Junio 11). *Proyecto Kelsen*. <http://derecho-internet.org/proyectos/documentacion/kelsen/teoria/informacion/ontologia>

De la Fuente, F. (2004). *Los sistemas de la información en la sociedad del conocimiento*. Madrid: ESIC Editorial.

Dechile.net. (2001). Consultada en enero 17, 2015. <http://etimologias.dechile.net/?Internet>

Decisión, No.2005-530 (Consejo Constitucional de Francia Diciembre 29, 2005). Consultada en febrero 3, 2015. <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000632146>

Der Jure.org. Rechtsinformationssysteme GmbH. (n.d.). Consultada en agosto 30, 2014. <http://dejure.org/>

Derecho Ecuador. (2013). Consultada en julio 15, 2014. <http://www.derechoecuador.com>

Derechoupr.com: tu Shortcut a la Investigación Jurídica. (2015). Consultada en agosto 5, 2014. <http://derechoupr.com>

Diritto.it. (n.d.). Consultada en marzo 20, 2014. www.diritto.it

DOE - The Differential Ontology Editor. (2008, Junio 24). Consultada en marzo 3, 2015. <http://www.eurecom.fr/~troncy/DOE/>

Doug Engelbart Institute. (2015, enero 25). *Curriculum Vitae*. Consultada en enero 25, 2015. <http://www.dougenelbart.org/about/cv.html>

Ecuador. Consejo de la Judicatura. (2011). Consultada en julio 14, 2014. <http://www.funcionjudicial.gob.ec/>

Ediciones Legales. (n.d.). Consultada en julio 15, 2014. <http://www.edicioneslegales.com.ec/>

El Peruano. Diario Oficial. (2010). Consultada en julio 26, 2014. www.elperuano.com.pe,

Especificación HTML 4.01. (n.d.). Consultada en marzo 7, 2015. ConClase.net: <http://html.conclase.net/w3c/html401-es/cover.html>

Espinel, Á., Guevara Bolaños, J. C., Flórez Fernández, H., Pérez Castillo, J. N., & Pinzón Nuñez, S. A. (2006). La web semántica y sus posibles aplicaciones en el ámbito universitario. *Revista Tecnura*(19), 121-131.

Estado Libre Asociado de Puerto Rico. (2013). <http://www2.pr.gov/Pages/servicefilter.aspx?cat=servicios&type=ciudadano>

EUR-Lex. (n.d.). Consultada en marzo 14, 2015. <http://eur-lex.europa.eu/homepage.html?locale=es>

Fallows , D., Lee, R., & Graham, M. (2004, Agosto 12). *Memo de datos en los motores de búsqueda*. Consultada en febrero 5, 2014. Pw Research Center: <http://translate.google.com/translate?hl=es&langpair=en%7Ces&u=http://ww>

w.pewinternet.org/Reports/2004/Data-Memo-on-Search-Engines/Data-Memo.aspx

Fernández-Coca, A. (1998). *Producción y diseño gráfico para la World Wide Web*. Barcelona: Paidós Ibérica.

Fierro, M. (n.d.). *Informática jurídica documental*.

FindLaw. (2015). Consultada en agosto 30, 2013. <http://public.findlaw.com/>

Findlaw. Suprema Corte. (2015). Consultada en octubre 3, 2013. www.ilrg.com/caselaw

Fix Fierro, H. (1990). *Informática y documentación jurídica*. México: Universidad Nacional Autónoma de México. UNAM.

Florez Fernández, H. (2006, Octubre 4). *Web Semántica aplicada al registro académico Institucional*. Consultada en mayo 23, 2015. http://www.konradlorenz.edu.co/images/stories/suma_digital_sistemas/2009_01/Informe_Investigacion_semantica.pdf

FOAF. (2015, Febrero 3). [http://en.wikipedia.org/w/index.php?title=FOAF_\(ontology\)&oldid=645417702](http://en.wikipedia.org/w/index.php?title=FOAF_(ontology)&oldid=645417702)

Frosini, V. (1978-1982). *Cibernética, diritto e societa*. (C. A.-T. Díaz, Trans.) Madrid: Tecnos.

García García, J. (1979). Problemática de la documentación jurídica. In *III curso de informática jurídica para juristas hispanoamericanos*.

- García Jiménez, A. (2004). Instrumentos de representación del conocimiento: tesauros versus ontologías. *Anales de la Documentación*, 79-95. <http://revistas.um.es/analesdoc/article/view/1691>
- García Marco, F. J. (2009). Perspectivas sobre el uso de la Web Semántica en el tratamiento de la información y documentación legislativa. *Revista Scire*, 15(1), 59-85. <http://ibersid.eu/ojs/index.php/scire/article/view/1769/1740>
- GDELT. Global Data Base of Events, Languages and Tones*. (013-2014). Consultada en diciembre 3, 2014. <http://gdeltproject.org/>
- Gillies, J., & Cailliau, R. (2000). *How the Web was Born: The Story of the World Wide Web*. New York: Oxford University Press.
- Giraldo Ángel, J. (1990). *Informática jurídica documental: manual para el sistema universitario de informática jurídica, SUIJ*. Bogotá: Temis.
- Girard, B. (2009). *The Google Way: How One Company is Revolutionizing Management as We Know it*. San Francisco: No Starch Press, Inc.
- Gobierno del Principado de Asturias*. (n.d.). Consultada en marzo 14, 2015. <https://sede.asturias.es/portal/site/Asturias/menuitem.048b5a85ccf2cf40a9be6aff100000f7/?vgnextoid=c0c756a575acd010VgnVCM100000bb030a0aRCRD&i18n.http.lang=es&calendarioPqBopa=true>
- Google compra YouTube «1.650 millones de dólares »*. (2006, Octubre 9). Consultada en febrero 5, 2014. <http://google.dirson.com/post/2885-google-compra-youtube/>

Google Empresa. (n.d.). *Nuestra historia en Profundidad*. Consultada en mayo 9, 2015. <https://www.google.com/about/company/history/?hl=es>

Google, condenado por colgar vídeos en su web sin permiso de los autores. (2011, Marzo 9). *El Espectador*. Consultada en marzo 5, 2014. <http://www.elespectador.com/tecnologia/google-condenado-colgar-videos-su-web-sin-permiso-de-lo-articulo-255768>

Gorostiza, I. (2009, Mayo 25). *Tutorial Cache Web: cómo gestionar el cacheo de nuestros contenidos*. hellogoogle.com: <http://www.hellogoogle.com/tutorial-cache-web/>

Gorrón A., L. M. (2009, Diciembre 01). *Informatica Jurídica Documental*. Consultada en julio 25, 2013. Slide Share: <http://www.slideshare.net/guest9ca8c4/informatica-juridica-documental>

Guzman Luna, J., López García, A. N., & Torres Pardo, D. (2006). Desarrollo de una ontología en el contexto de la web semántica a partir de un tesoro documental tradicional. *Revista Interamericana de Bibliotecología*, 29(2), 79-94.

Harley, B. (2010, Marzo 31). Semantic Lawyering: How the Semantic Web Will Transform the Practice of Law (Part 1). *Columbia Science and Technology Law Review*, Parte 1. Semantic Lawyering: How the Semantic Web Will Transform the Practice of Law: <http://stlr.org/2010/03/31/semantic-lawyering-how-the-semantic-web-will-transform-the-practice-of-law-part-1/>

Hendler, J. (2009). Web 3.0 emerging. *IEEE Journals & Magazines*, 42(1), 111- 113 . doi:10.1109/MC.2009.30

Hernández Gómez, J. R. (2007). *Antropología jurídica Marcel Mauss el derecho del potlatch*. Bogotá.

Hernández Marin, R. (2009). *Sobre ontología jurídica e interpretación del derecho*. Alicante: Biblioteca Miguel de Cervantes.

Hipertexto. (2015, enero 23). Wikipedia, La enciclopedia libre:
<http://es.wikipedia.org/wiki/Hipertexto>

Historia de Internet. (2015, Febrero 14). Consultada en febrero 5, 2013. Wikipedia:
http://es.wikipedia.org/w/index.php?title=Historia_de_Internet&oldid=80016630.

Hohfeld, W. (1969). *Concetti giuridici fondamentali*. Turin: Giulio Einaudi.

HTML 4 para dummies. (1997). Panama: IDG BooksWorldwide.

<http://supcourt.ntis.gov/>. (n.d.).

Imprenta Nacional de Colombia. (2015). Consultada en agosto 30, 2014.
<http://www.imprenta.gov.co/portal/page/portal/inicio>

Índice Fundamental do Direito. (2014, Septiembre 9). Consultada en julio 10, 2014.
<http://www.realjus.com.br/dji/>

Inforlex. Sistema de índices jurídicos. (1993-2013). Consultada en agosto 26, 2014.
<http://www.inforlex.com.uy/>

Informática Jurídica Documentaria. (n.d.). In *[Informática instrumento del derecho]*.
<http://biblio.juridicas.unam.mx/libros/4/1941/7.pdf>

Informática Jurídica.com. (2015, Marzo 28). Consultada en julio 7, 2014.

<http://www.informatica-juridica.com>

Informática Legal. (2011). Consultada en julio 7, 2014.

<http://www.informaticalegal.com.ar/legislacion-informatica/>

InfoSynthesis Puts U.S. Supreme Court Decisions. (1998, Noviembre). Consultada

en septiembre 23, 2013 <http://connection.ebscohost.com/c/articles/1296630/>

[infosynthesis-puts-u-s-supreme-court-%20decisions-cd](http://connection.ebscohost.com/c/articles/1296630/infosynthesis-puts-u-s-supreme-court-%20decisions-cd).

Institut Canadien d'Information Juridique. CanLII. (n.d.). Consultada en marzo 14,

2015. <http://www.canlii.org/>

Instituto de Investigaciones Jurídicas. (1995). Consultada en julio 20, 2014.

<http://www.juridicas.unam.mx/>

Instituto Mexicano de Estrategias S.C. (n.d.). Consultada en julio 20, 2014.

<http://www.siemexico.mx/>

International Network for Economic, Social & Cultural Rights. ESCR-NET. (n.d.).

Consultada en julio 5, 2014. <http://www.escr-net.org/>

Internet Legal Research Group. (2015, Enero 20). Consultada en octubre 3, 2013.

<http://www.ilrg.com/caselaw/>

Italgireweb. (n.d.). Consultada en enero 27, 2014. www.italgiure.giustizia.it

Italia. Gazzetta Ufficiale. (n.d.). Consultada en marzo 6, 2014.

<http://www.gazzettaufficiale.it/>

Ittig. Institute of Legal Information Theory and Techniques. (n.d.). Consultada en

mayo 7, 2014. <http://www.ittig.cnr.it/>

Joyanes Aguilar, L. (2009). *Redes sociales: la nueva revolución de la era de facebook*. España: Joxman Editores Multimedia.

Joyanes Aguilar, L. (2013). *Análisis de grandes volúmenes de datos en organizaciones*. México: Alfaomega Grupo Editor.

Junta General del Principado de Asturias. (n.d.). Consultada en marzo 14, 2015.

Junta General del Principado de Asturias: <http://www.jgpa.es/>

Juris. (n.d.). Consultada en septiembre 1, 2014. *Juris. Recht Portal*:

<http://www.juris.de/jportal/nav/kostenfreieinhalte/infokostenfreieinhalte.jsp>

Juristisches Internetprojekt Saarbrücken. (n.d.). Consultada en agosto 29, 2014.

<http://www.jura.uni-saarland.de/deutschland/>

Juriversia. (n.d.). Consultada en julio 21, 2014. <http://www.juriversia.com/>

Mexico/index.aspx

Jurix. The Foundation for Legal Knowledge Based Systems. . (2015). <http://jurix.nl/>

Kah, R. (2014, Junio 17). *Wikipedia, La enciclopedia libre*. Consultada en marzo 7,

2015. http://es.wikipedia.org/w/index.php?title=Robert_Kahn&oldid=75085395.

Kelsen, H. (2000). *Teoría pura del derecho* (4 ed.). México: Colofón.

La Ley Online. (2011). Consultada en julio 7, 2014. <http://www.laleyonline.com.ar>

La Rama Judicial de Puerto Rico. (n.d.). Consultada en agosto 5, 2014.

<http://www.ramajudicial.pr/>

Lancho Pedrera, F. (2003). Los sistemas expertos en el derecho. *Anuario de la Facultad de Derecho de la Universidad de la Rioja*, XXI, 629-536.

Larousse. (n.d.). Consultada en marzo 5, 2015.
http://www.larousse.fr/encyclopedie/article/LIntelligence_Artificielle/11011577

LawGuru. (1997). Consultada en agosto 02, 2013. LawGuru:
<http://www.lawguru.com/>

Lawrence, S., & Giles, L. (n.d.). *Metrics.com*. <http://www.wwwmetrics.com/>

Le Boulengue de Galard, M. (n.d.). El tratamiento electrónico de la documentación jurídica. *Anuario de la Facultad de Derecho*(4).

Le Conseil d'État et la Juridiction Administrative. (n.d.). Consultada en marzo 13, 2015. Le Conseil d'État et la Juridiction Administrative: <http://www.conseil-etat.fr/Decisions-Avis-Publications/Decisions/ArianeWeb>

Le Portail du Pouvoir Judiciaire de Belgique. (n.d.). Consultada en marzo 13, 2015. Juridat: <http://www.juridat.be>

Legal Research Services. (2009). Consultada en octubre 3, 2013.
<http://www.lawresearchservices.com>

Legifrance. (n.d.). Consultada en marzo 13, 2015. Legifrance:
<http://www.legifrance.gouv.fr/>

Legis Colombia. (n.d.). Consultada en marzo 13, 2015. <http://www2.legis.com.co/>

Legis Law. (1999). Consultada en julio 10, 2014. www.legislaw.com.ar

Levering, R., & Moskowitz., M. (n.d.). *100 best companies to work for*. Consultada en enero 12, 2012. Great Place to Work: <http://www.greatplacetowork.com/best-companies/100-best-companies-to-work-for>

Lex Nova. (n.d.). Consultada en marzo 14, 2015. Lex Nova: <http://www.lexnova.es/>

Lexbase. (n.d.). Consultada en agosto 28, 2014. <http://www.lexbase.com/>

Lexis. (2015). Consultada en julio 15, 2014. <http://goo.gl/zzmG9B>

Lexis. Inteligencia Jurídica. (1998). Consultada en julio 15, 2014. <http://www.lexis.com.ec>

LEXISNEXIS ACADEMIC. (2015). Consultada en junio 22, 2015. *LexisNexis Academic*: [http://www.lexisnexus.com/en-us/products/lexisnexus-academic.page./](http://www.lexisnexus.com/en-us/products/lexisnexus-academic.page/)

Lexml.es. (n.d.). Consultada en marzo 13, 2015. Lexml.es: <http://www.uv.es/lexml/>

LexUM. (2015). Consultada en marzo 14, 2015. LexUM: <https://lexum.com/fr/>

Leyes.info. Información Económica y Jurídica de Colombia. (2015). Consultada en agosto 28, 2014 http://www.leyex.info/index_descripcion.php

López Muñoz-Goñi, M. (1984). *La informática jurídica documental*. Madrid: Díaz de Santos.

López-Muñoz Goñi, M. (1979). *Curso de informática jurídica*. Madrid: Fundesco.

Losano, M. G. (1969). *Giuscibernética Macchine Ciberneticici nel Diritto*. Torino: Einaudi.

Losano, M. G. (1982). *Introducción a la Informática Jurídica*. España: Universidad de la Palma Mallorca. Facultad de Derecho.

L'Union Européenne. (n.d.). Consultada en marzo 15, 2015. L'Union Européenne:
http://europa.eu/index_fr.htm

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011, Mayo). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute:
<http://goo.gl/3rwzGX>

Manzanares, H., & Nectoux, P. (1988). *La informática al servicio del jurista*. Bogotá: Legis.

Marco, M. A. (n.d.). *L'informatica giuridica ed il sistema Italgiure*. Consultada en enero 27, 2014.
<http://www.corsomagistratitributari.unimi.it/STRUTTURA/DOTTRINA%20E%20RELAZIONI/VARIA/Approfondimenti/40%20Alma/A-40%20Alma.pdf>

Márquez Solís, S. (2007). *La web semántica y servicios web semánticos*. Universidad Oberta de Cataluña.

Martín, H., & Priscila, L. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science Magazine*, 332(6025), 60.65.

Martino, A. A. (1985). *La informática jurídica hoy*. Buenos Aires: Ediciones Depalma.

Mazet, G. (1983). Accés en différentes langages a un Corpus Documentaire
 Problematique et methodologie. In *III Congresso Internazionale su
 L'informática Giuridica e la Comunita Nazionale ed Internazionali*. Roma.

Mazzega, P., Boucier, D., Bourguine, P., & Nadah, N. (n.d.). *Ontologie juridique et
 systemes complexes*. Consultada en febrero 4, 2015.
<http://www.gemas.fr/dphan/rochebrune09/papiers/BoucierDanielet.pdf>

México. Cámara de Diputados. (2006). Consultada en julio 21, 2014.
<http://www.diputados.gob.mx/LeyesBiblio/index.htm>

México. Corte Suprema de Justicia. (n.d.). Consultada en julio 21, 2014.
<https://www.scjn.gob.mx/Paginas/Inicio.aspx>

Microjuris.com. *Inteligencia Jurídica*. (1989-2015). Consultada en agosto 5, 2014.
<http://pr.microjuris.com/products-MJSuperior.jsp>

Moreno Jiménez, P. M. (2005). *Estrategias y mecanismos de búsqueda en la Web
 Invisible*. Bogotá: Universidad Militar Nueva Granada.
<http://www.umng.edu.co/www/resources/web%20invisible.pdf>

Moriello, S. A. (2013, Septiembre 13). *Investigación: Sistemas complejos*.
 Consultada en marzo 2, 2015. <http://www.cs.us.es/~fsancho/?p=sistemas-complejos-2>.

Multilegis. *Conexión a la Información Jurídica*. (2011). Consultada en agosto 28,
 2014. <http://www.multilegis.com/colombia.html>

Muñoz López, M. (2010). Humanistas, textos e Internet: una difícil relación. In J. M.
 Maestre Maestre, A. Prieto, J. P. Barea, & L. Charlo Brea, *Humanismo y*

pervivencia del mundo clásico homenaje al profesor Antonio Prieto (Vol. 4, pp. 1993-2000). Alcañiz (Madrid): Consejo Superior de Investigaciones Científicas. CSIC. Consultada en marzo 5, 2014. <https://books.google.com.co/books?id=o81r4gQv-ewC&pg=PA1999&lpg=PA1999&dq=Humanistas,+textos+e+Internet:+Una+dif%C3%ADcil+relaci%C3%B3n&source=bl&ots=wXHOR3Cw5x&sig=UsIONHm2CF0UawKqOM-6YiVXzR4&hl=es&sa=X&ei=pAfhVLCUI4eqgwSbyINQ&ved=0CCQQ6AEwAw#v=onepage&q>

Norma Jurídica. Institute of European and Comparative Private Law. (2014, Junio 4). <http://civil.udg.es/NORMACIVIL/>

Noti.net. (2015). Consultada en agosto 28, 2014. <http://www.notinet.com.co/>

Noticias Jurídicas. (n.d.). <http://noticias.juridicas.com/>

Oasis LegalXML. (2005, Mayo). <https://www.oasis-open.org/committees/download.php/12785/OASIS%20LegalXML%20Brochure.pdf>

Oficina de Servicios Legislativos de la Asamblea Legislativa de Puerto Rico. (n.d.). Prontuario: <http://www.oslpr.org/prontuario/>

Page Rank. (2015, Enero 27). Consultada en febrero 15, 2015. <http://es.wikipedia.org/w/index.php?title=PageRank&oldid=79638444>.

Page Rank, la idea de Larry que cambió el mundo. (2011, Abril 2). Consultada en febrero 5, 2014. Enter.co: <http://www.enter.co/cultura-digital/redes-sociales/pagerank-la-idea-de-larry-que-cambio-al-mundo/>

Panamá. Comisión Nacional de Reformas Electorales. (n.d.). Consultada en agosto 26, 2014. <http://www.tribunal-electoral.gob.pa/html/>

Panamá. Procuraduría de la Administración. Infojurídica. (2012). Consultada en agosto 26, 2014. <http://goo.gl/JV7UTl>

Parlament de les Illes Balears. (n.d.). Consultada en marzo 14, 2015. Parlament de les Illes Balears: <http://www.parlamentib.es/Home/Default.aspx>

Parlamento de Cantabria. (n.d.). Consultada en marzo 14, 2015. Parlamento de Cantabria: <http://www.parlamento-cantabria.es/>

Parlamento Italiano. Camera dei Deputati. (n.d.). Consultada en enero 20, 2014. <http://goo.gl/dND3kQ>

Parlamento Vasco. (n.d.). Consultada en marzo 14, 2015. <http://www.parlamentovasco.euskolegebiltzarra.org/es>

Parlement Européen. (2014, enero 18). Archives: <http://www.europarl.europa.eu/parlArchives/comArch/comDocSearch.do;jsessionid=B91E98AD4078138EE6655FB2382A475E.node1?language=fr>

Peiron Guardia, M., & Sánchez Carracedo, F. (2002). *Los circuitos lógicos combinacionales*. Barcelona: UOC.

Pérez Luño, A. E. (1996). *Manual de informática y derecho*. Barcelona: Ariel.

Perú. Congreso de la República. (n.d.). Consultada en julio 26, 2014. <http://www.congreso.gob.pe/>

Portal Europeo de e-Justicia. (n.d.). Consultada en agosto 5, 2014. 1 Portal Europeo de e-Justicia: <https://e-justice.europa.eu/home.do>

Portal Jurídico DMS. (n.d.). <http://www.portaldms.com/index.html>

Portal Jurídico El Derecho. (n.d.). Consultada en marzo 14, 2015. Portal Jurídico El Derecho: <http://www.elderecho.com/>

PortalDerecho. Iustel. (2001). <http://www.iustel.com/>

Portal Jurídico. (2013, Agosto 1). <http://www.portaldms.com/index.html>

Rama Judicial de Colombia. (2015). Estadísticas Judiciales Enero 1 a Diciembre 31 de 2014: <http://www.ramajudicial.gov.co/web/estadisticas-judiciales/movimiento-de-procesos-del-1-de-enero-a-31-de-diciembre-2014>

Ramírez Cleves, G. (2010). *Los blogs jurídicos y la web 2.0 para la difusión y la enseñanza del derecho.* Bogotá: Universidad Extenado de Colombia.

Salazar Cano, E. (1978). *Informática jurídica. Instrucción programada y documentación automática.* Caracas: Universidad de Carabobo.

Sánchez Mazas, M. (n.d.). *Panorama general de la Informática Jurídica.* Fundación para el Desarrollo de la Función Social de las Comunicaciones. FUNDESCO.

Sartor, G. (2006). Il diritto nella societ dell'informazionne. In G. Sartor, *Informatica Giuridica.* Consultada en abril 30, 2015. http://www2.cirsfid.unibo.it/didattica/upload/176_GS2005LIGTestoCap1.pdf

Schweghofer, E., & Liebwald, D. (2008, Junio). Advanced lexical ontologies and hybrid knowledge bases systems: first steps to a dynamic legal electronic commentary. *Artificial Intelligence and Law*, 15(2), 103-115.

Senat. (n.d.). Consultada en marzo 13, 2015. Senat: <http://www.senat.fr/>

Serrano Pirela, A. (1975). *Computadoras y derecho: una introducción a la informática jurídica*. Maracaibo: Universidad de Zulia. Facultad de Derecho, Centro de Estudios de Filosofía del Derecho.

Service Public Fédéral Justice. (2013). Consultada en marzo 13, 2015. Service Public Fédéral Justice: <http://justice.belgium.be/fr/>

Sherman, C., & Price, G. (2007). *The invisible web. Uncovering information sources search engines can't see*. New Jersey: Information Today.

Siegel, D. (2009). *The power of the Semantic Web to transform your business group*. Penguin Group.

Simple HTML Ontology Extensions. SHOE. (n.d.). <http://www.cs.umd.edu/projects/plus/SHOE/>

Siri, L. (2000). *Internet: búsquedas y buscadores*. (Vol. 2). Bogotá: Norma.

Sistema Argentino de Información Jurídica. INFOJUS. (n.d.). Consultada en julio 2, 2014. <http://www.infojus.gob.ar/>

Sistema Costarricense de Información Jurídica. SCIJ. (n.d.). Consultada en julio 27, 2014. Sistema Costarricense de Información Jurídica, SCIJ

Sistema de Archivo de Documentos de las Naciones Unidas. (2004). <http://documents.un.org/>

Sistema de Cooperación Jurídica. SCJ. (n.d.). Consultada en mayo 29, 2014. <http://www.scjuridica.org.ve/portal/php/001.php>

Sistema Peruano de Información Jurídica. SPIJ. (n.d.). Consultada en julio 22, 2014 <http://spij.minjus.gob.pe>

Stanford Center for Biomedical Informatics Research. (2015). *Protégé*. Consultada en febrero 3, 2015. <http://protege.stanford.edu/>

Start making sense:. (2008, Abril 9). *The Economist: Big and small companies are getting into the business of building an intelligent web of linked data.* <http://www.economist.com/node/11002939>

Suñé Llinás, E. (1986). Introducción a la informática jurídica y al derecho de la informática. *Revista de la Facultad de Derecho de la Universidad Complutense de Madrid*(12).

Suñé Llinás, E. (1988). Los tesauros en bases de datos de usuario. In *Actas del congreso de informática y derecho* (pp. 1-27). Roma.

Suñe Llinás, E. (1994). *Informática práctica para juristas y profesionales del mundo de las letras*. Madrid: Servicio de Publicaciones, Facultad de Derecho, Universidad Complutense y Beramar.

Superley. Información Legal al Día. (n.d.). Consultada en julio 15, 2014. <http://www.superley.ec/super/>

Suprema Corte de Justicia de la Nación. (n.d.). Tesauro Jurídico: https://www.scjn.gob.mx/Sistema_Bibliotecario/Comunicados/Paginas/tesauro.aspx

Suprema Corte de Justicia de la Nación. (n.d.). Consultada en agosto 5, 2014. Biblioteca Digital y Sistema Bibliotecario: https://www.scjn.gob.mx/Sistema_Bibliotecario/Paginas/BasesD_esp.aspx

Synthèse du marché Big Data. (2013). *Big Data*, (pp. 1-24). París.

http://corpevents.fr/bigdata14/docs/Dossier_de_synthese_BIG_DATA_PARI_S_2013.pdf

Szulman, S., Hernández, N., Teguiak, V., Charlet, J., Nadah, N., Aussenac-Gilles, N., . . . Delahousse, J. (2010). DAFOE : une plateforme multiméthodes et multimodèles pour. *Congres Francophone AFRIF-AFIA Reconnaissance des Formes et Intelligence Artificielle*. l'Université de Caen Basse-Normandie. Consultada en marzo 28, 2015. <ftp://ftp.irit.fr/IRIT/IC3/demoRFIA.pdf>

Temprano Sánchez, A. (2008). *Diseño y desarrollo de un software para la creación de webquest*. Madrid: Bubok.

The Bodleian Law Library. (2014). Consultada en agosto 8, 2014. Bases: <http://www.bodleian.ox.ac.uk/law>

The Estrella project website. (n.d.). <http://www.estrellaproject.org/>

The European e-Justice Portal. (2014). Consultada en marzo 13, 2015. The European e-Justice Portal: <https://e-justice.europa.eu/home.do?plang=en&action=home>

The FedWorld a program of the United States Departament of Commerce. (n.d.). Consultada en julio 2013, 2013. Supreme Court Decisions: <http://supcourt.ntis.gov/>

The German Law Archive. (2006). Consultada en agosto 29, 2014. <http://www.iuscomp.org/gla>

The Parliament of the United Kingdom of Great Britain and Northern Ireland. (n.d.).

Consultada en marzo 14, 2015. The Parliament of the United Kingdom of Great Britain and Northern Ireland: <http://www.parliament.uk/>

The Supreme Court of United Kingdom. (2015). Consultada en marzo 14, 2015.

Decided cases: <https://www.supremecourt.uk/index.html>

Tirant Lo Blanch. (n.d.). Consultada en marzo 14, 2015. Tirant Lo Blanch:

<http://www.tirantonline.com/tol/>

Todoelderecho.com. (2014). Consultada en julio 6, 2014.

<http://bases.todoelderecho.com>

Torres Andrade, C. A. (2008). Agentes y web semántica: elementos claves para el futuro del e-learning. *Revista Tecnura*, 11(22), 111-123.

Torres, A. (2009, Mayo 3). *WEB 2.0 Qué es y que nos espera?* Consultada en marzo

5, 2012. WEB 2.0+TIC hacia donde vamos!: <https://aulaticujap.wordpress.com/unidad-i/about/>

Tribunal des Conflicts. (n.d.). Consultada en marzo 13, 2015. Decisions:

http://www.tribunal-conflicts.fr/decisions_2014.html

Tribunal de Justicia de la Unión Europea. (2014, enero 18).

http://curia.europa.eu/jcms/jcms/j_6/

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 49, 433-460.

<http://www.csee.umbc.edu/courses/471/papers/turing.pdf>

UK. Legislation. (n.d.). Consultada en agosto 5, 2014. UK. Legislation:

<http://www.legislation.gov.uk/>

Ulpiano. Primera fuente de referencia jurídica de Venezuela. (n.d.). Consultada en junio 5, 2014. <http://www.ulpiano.org.ve/>

Union Européenne. (2014, enero 18). *Site web officiel de l'Union européenne.* Législation de l'UE: http://europa.eu/eu-law/index_fr.htm

Universidad de Chile. Biblioteca Central Facultad de Derecho. (n.d.). Guía de uso Lexis Nexis. Chile. Consultada en mayo 29, 2015. [file:///D:/Users/hiomara.vargas/Downloads/base%20de%20datos%20lexis%20onexis%20\(1\).pdf](file:///D:/Users/hiomara.vargas/Downloads/base%20de%20datos%20lexis%20onexis%20(1).pdf)

Università di Bologna. Centro Interdipartimentale di Ricerca in Storia del Diritto, Filosofia e Sociologia del Diritto e Informatica Giuridica. CIRSFID. (2015). Consultada en marzo 20, 2014. <http://www.cirsfid.unibo.it/>

Université Laval. Faculte des Sciences Sociales. Département des Relations Industrielles. (2015). Consultada en marzo 14, 2015. Documentos: <http://www.rlt.ulaval.ca/documents/accueil.asp>

University of Bologna. (2015). *Akoma Ntoso: XML for parliamentary, legislative and judiciary documents.* <http://www.akomantoso.org/>

University of London. School of Oriental and African Studies. (2007). R Consultada en agosto 25, 2014. Bases de datos: <http://www.soas.ac.uk/library/subjects/law/databases/lexisnexis-butterworths/>

Uruguay. Contaduría General de la Nación. (n.d.). Consultada en agosto 26, 2014. <http://goo.gl/vnKr8L>

Vargas Osorno, T. G. (1993). *Manual de introducción a la informática jurídica*.
Bogotá: Universidad Externado de Colombia.

Venezuela. *Tribunal Supremo de Justicia*. (2014). Consultada en junio 5, 2014.
<http://www.tsj.gob.ve/>

vLex. *Argentina*. (2014). Consultada en julio 7, 2014.
<http://ar.vlex.com/jurisdictions/AR>

Westlaw UK. (2015). Consultada en marzo 14, 2015. Westlaw UK:
<http://legalresearch.westlaw.co.uk/>

World Wide Web Consortium. (n.d.). *Guía breve de Web Semántica*. Consultada en
noviembre 8, 2014. [http:// www.w3c.es/Divulgacion/Guiasbreves](http://www.w3c.es/Divulgacion/Guiasbreves).

Yonck, R. (2011, Enero 17). *What is a Milestone in Artificial Intelligence?*
Consultada en febrero 2, 2015. World Future Society: <http://intelligent-future.com>

BIBLIOGRAPHY

- Alvite Díez, M. L. (2009, Enero-Junio). Las bases de datos jurídicas y el uso del lenguaje XML en España. *Scire*, 15(1), 2. <http://www.iberid.eu/ojs/index.php/scire/article/view/1768/1739>
- Berners-Lee, T., Hendler, J., & Lasilla, O. (2001). The Semantic Web: a new form of web content tha is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, 284(5), 34-43.
- Casanovas, P. (2012, Enero-Junio). Algunas líneas de investigación en gestión del conocimiento jurídico: web semántica, ODR y derecho racional. *Revista Scire*, 18(1), 15-28. <http://iberid.eu/ojs/index.php/scire/article/viewFile/3943/3672>
- Castells, P. (2003). La Web Semántica. In C. Bravo Santos, & Á. Redondo Duque, Sistema interactivos y colaborativos en la Web (p. 280). Cuenca: Universidad de Castilla-La Mancha. <https://goo.gl/6CXQrO>
- Frosini, V. (1978-1982). *Cibernética, diritto e societa*. (C. A.-T. Díaz, Trans.) Madrid: Tecnos.
- Joyanes Aguilar, L. (2013). *Análisis de grandes volúmenes de datos en organizaciones*. México: Alfaomega Grupo Editor.
- López Muñiz-Goñi, M. (1984). *La informática jurídica documental*. Madrid: Díaz de Santos.
- Losano, M. G. (1982). *Introducción a la Informática Jurídica*. España: Universidad de la Palma Mallorca. Facultad de Derecho.
- Mazzege, P., Boucier, D., Bourguine, P., & Nadah, N. (n.d.). *Ontologie juridique et systemes complexes*. Consultada en febrero 4, 2015. <http://www.gemas.fr/dphan/rochebrune09/papiers/BoucierDanielet.pdf>
- Pérez Luño, A. E. (1996). *Manual de informática y derecho*. Barcelona: Ariel.
- Sánchez Mazas, M. (n.d.). *Panorama general de la Informática Jurídica*. Fundación para el Desarrollo de la Función Social de las Comunicaciones. FUNDESCO.
- Siri, L. (2000). *Internet: búsquedas y buscadores*. (Vol. 2). Bogotá: Norma.
- Suñe Llinás, E. (1994). *Informática práctica para juristas y profesionales del mundo de las letras*. Madrid: Servicio de Publicaciones, Facultad de Derecho, Universidad Complutense y Beramar.

ABSTRACT

LEGAL DATABASES: PAST, PRESENT AND FUTURE

The evolution of legal databases is evident. One of the main causes is that law is making constant changes in addition to the emergence of new technological tools and services available to users, allowing going beyond the simple organization of information.

At the beginning, documentary legal information arose from the urgent need to organize legal information to facilitate the work of professionals in law. With its creation, tree structures formed by descriptors appeared which as a whole constitute the Thesaurus. Thanks to these, we can see the organization of legal reasoning and the *raison d'être* of the structure in which legal documents are drafted by lawmakers, judges or various relevant officials, to publish documents with binding content.

In the past, documentary legal information was able to organize the information so that the user of legal database could locate, in a short time, information in accordance with his/her requirement and object of the search.

In some of the legal databases from the past, it was possible to examine the relevant legal views of the documents studied, since the analyst carried out identification tasks and a synthesis of the information provided.

One of the main aspects that has evolved in the documentary legal information refers to the storage capacity and the way to locate it. At the beginning, storing documents in a database was performed using a floppy disk or CD.

With Internet, Web 1.0 is born allowing users to access information online. Internet, in turn, provides greater storage capacity with the existence of servers that offer storage services for large amounts of information.

Similarly, Web 2.0 emerges, the so-called social or participatory web, in which users besides consulting information can contribute to the construction of information. This dynamic makes online platforms offering legal databases, to deliver new services in which the user can play an active role in the creation of relevant information.

Despite this, one of the critics in relation to the use of Web 2.0 is the reliability of the information that rests on the Internet, since the participation prevents from exercising control over the quality of information provided by the user on line, which can create instability and affect the work of the professional.

With the unusual use of the Internet, the phenomenon of information overload arises. This phenomenon of knowledge and normative expansion and the tendency of law to regulate every assumption made, leads to the inability to identify within the mass of information, the one relevant to the user who undertakes the search.

In the era of the information society, the current problem does not only consist of knowing how to store and organize information but how to get timely and relevant search results when you start searching. Consequently, current information overload

problems is the inability to analyze large amounts of data in a logical and reasonable manner.

Thanks to the emergence of major web browsers, a fundamental change for documentary legal information is created, since there are new forms of information search and particularly, new ways to retrieve it, thus the search is optimized and the implementation of these techniques in legal databases leads us to obtain timely results.

With the arrival of Web 1.0 and Web 2.0., and especially with the emergence of specialized search engines, documentary legal information at present, the need to evolve and adapt to new technologies is seen. At present, the goal of a legal data bank goes beyond retrieving and organizing information; it seeks to concatenate, store and provide more complete and relevant results.

In the future, the purpose of documentary legal information will vary in relation to legal information today, since the purpose will not only consist of finding the best way of organizing information but to make data do the talking and get from rational and logical answers to structure data contents simulating human legal rationing.

Consequently, the main purpose is not only to find a document but also to obtain constructed and structured knowledge because of the search. It is important to note also that the future to which we refer to is immediate. Changes already are on the way, so the concept of legal information has evolved, thanks to the emergence of the

Semantic Web, the implementation of artificial intelligence techniques, particularly through the creation and implementation of legal ontologies.

There are thus two phenomena that have started to make changes on documentary legal information, these are the emergence of artificial intelligence applied to law and the creation and implementation of legal ontologies.

As to the first phenomenon, related to artificial intelligence, it affects legal information document in two ways: first, through the creation of expert systems capable of simulating the human legal reasoning to solve a case. Second, the creation of legal ontologies enabling the creation of intelligent legal databases, which manage that their data do the talking by throwing information with semantic content. The latter concerns us, since the structure of legal documents and the meaning of its contents is analyzed thanks to the study of the structure.

Therefore, legal ontologies play a fundamental role in the construction of knowledge, within a smart legal database. Ontologies represent and organize the contents of the information, using category hierarchies and the relationship between tree structures, managing and organizing not only a database but also thesauri interconnection.

Ontology, in turn, can apply the use of statistics to create odds ratios between concepts, bring descriptors closer and interconnect information, which would be related if a human analyst were to perform the task.

We conclude that the progress of documentary legal information, since its inception can be easily ascertained; current databases and those of the near future will not be satisfied with the simple goal of organizing information to facilitate the search as if it were a card index file or a library. These databases aim at achieving interconnection of documents and concepts contained therein, deliver results to the user by connecting contents with semantic concepts, matching them to facilitate their application.

We are facing major challenges. Legal artificial intelligence may contribute to the activity of the legal operator (magistrates, judges, lawyers, notaries, etc.). Semantic and rational results can facilitate decision-making and compliance with law by responding in real-time to changes that occur within the legal system.